MAT 1741		**1. Nombres complexes**
Mike Newman	Université d'Ottawa	automne 2008

IMAGINAIRE

Vous connaissez probablement comment résoudre $x^2 - 5x + 6 = 0$: on peut factoriser pour donner $x = 2, 3$. Les solutions de $x^2 - 2 = 0$? On les connait, ce sont $\pm\sqrt{2}$.

Mais qu'est-ce, exactement, $\sqrt{2}$? Ma calculatrice me dit $1,414213562$, mais c'est clairement un mensonge, car $\sqrt{2}$ n'est pas rationnel. En fait, la définition de $\sqrt{2}$ est précisement que c'est une racine de $x^2 - 2 = 0$. C'est un chiffre déterminé par une propriété.

C'est dans ce veine qu'on veut poser la question suivante: quelles sont les solutions de $x^2 + 1 = 0$? Les solutions ne sont certainement pas des nombres réels, car tout réel au carré devient non-négatif. On *définit* la valeur i comme étant une racine de $x^2+1 = 0$. On dit parfois que "$i = \sqrt{-1}$", mais c'est plus correct de dire simplement que c'est une valeur qui est racine de $x^2 + 1 = 0$.[1]

Question: On a l'habitude de penser que les quadratiques ont deux racines. Quelle est l'autre racine de $x^2 + 1 = 0$?

COMPLEXES

Étant donné i, on a les nombres complexes.

Définition 1.

$$\mathbb{C} = \{a + bi | a, b \in \mathbb{R}\}$$

En mots, un nombre complexe z peut s'écrire de la forme $a + bi$, où a et b sont des nombres réels et i est une racine de $x^2 + 1 = 0$.

Soit un nombre complexe $z = a + bi$. On dit que $\mathrm{Re}(z) = \mathrm{Re}(a + bi) = a$ est la partie réelle de z, et que $\mathrm{Im}(z) = \mathrm{Im}(a + bi) = b$ est la partie imaginaire. On peut donc placer z dans le *plan complexe*. C'est-à-dire, on a un axe "réelle" et un axe "imaginaire". On a donc une interprétation géométrique de z. On verra un tel dessin en classe…

NB: $\mathrm{Im}(a + bi) \neq bi$

[1] La valeur i est parfois dit "imaginaire". Mais c'est un nom mal-choisi. Par exemple, en physique les champs magnétiques et électriques sont fortement reliés. La description de cette interaction ce fait naturellement à l'aide de i. Une exemple de cette intéraction est la lumière: on dirait pas que la lumière est "imaginaire"!

ARITHMÉTIQUE

L'arithmétique des complexes est fondée sur les réels. On peut penser que ce sont des polynomes avec variable "i", avec la réduction $i^2 = 1$.

$$\begin{aligned}(a+bi) \pm (c+di) &= (a \pm c) + (b \pm d)i \\ (a+bi)(c+di) &= ac + adi + bic + bidi \\ &= ac + (ad+bc)i + bdi^2 \\ &= (ac - bd) + (ad+bc)i\end{aligned}$$

On a aussi deux autre opérations: la *conjugué complexe*

- $\overline{a+bi} = a - bi$

et la *valeur absolue*

- $|a+bi| = \sqrt{a^2+b^2}$.

Question: On note que $z = a + bi$ est réel si et seulement si $b = 0$. Est-ce que l'arithmétique ci-haut correspond à l'arithmétique réel si on met $b = d = 0$?

La conjuguée complexe à une interprétation géométrique: réflection dans l'axe réel. La valeur absolue correspond à la distance de l'origine. On verra en classe...

On a quelques propriétés de la valeur absolue et de la conjugué complexe.

- $|z| \geq 0$
- $|z| = 0 \iff z = 0$
- $z = \overline{z} \iff z \in \mathbb{R}$
- $|z| = |\overline{z}|$
- $\overline{z+w} = \overline{z} + \overline{w}$
- $\overline{(\overline{z})} = z$

Question: Pour chaque propriété, donner une interprétation géométrique. Qu'arrive-t-il si $z = a + 0i$ est réelle?

Question: Est-ce vrai que $|z+w| = |z| + |w|$? (indice: la valeur absolue equivaut à la distance de l'origine...)

Il reste une chose: comment diviser? En fait, puisque

$$\frac{z}{w} = z\frac{1}{w}$$

la vraie question c'est comment trouver l'inverse. On note que

$$\frac{1}{c+di} = \frac{1}{c+di}\frac{c-di}{c-di} = \frac{c-di}{c^2+d^2} = \frac{c-di}{|c+di|^2} = \frac{c}{c^2+d^2} - \frac{d}{c^2+d^2}i$$

On voit alors que

$$\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{c^2+d^2}.$$

On laisse la multiplication au numérateur comme exercice!

En générale, on a que $\frac{1}{z} = \frac{\overline{z}}{|z|^2}$, ou plutôt

- $z\overline{z} = |z|^2$

POLYNOMES

Avec les complexes, on trouve les racines de *n'importe quel* polynome quadratique. Par exemple, $x^2 + 2x + 3 = 0$. La formule quadratique donne que les racines sont $1 \pm \sqrt{2}i$. On vous laisse les calculs comme exercice en arithmétique complexe. Noter que les racines sont conjugués.

En générale, si $a, b, c, \in \mathbb{R}$, alors le polynome $p(x) = ax^2 + bx + c = 0$ possède deux racines. Mettons $\Delta = b^2 - 4ac$. On a trois cas:

$\Delta > 0$	deux racines réeles	$p(x) = a(x - \alpha_1)(x - \beta)$
$\Delta = 0$	une racine rélle répétée	$p(x) = a(x - \alpha)^2$
$\Delta < 0$	deux racines complexes conjuguées	$p(x) = a(x - \gamma)(x - \overline{\gamma})$

Exercice: : Déterminer un polynome au coefficients réels ayant comme racine $1+i$.

L'applicabilité dépasse les quadratiques. À l'aide des complexes, *tout* polynome peut se factoriser. C'est la théorème fondamentale de l'algèbre.

Théorème 2. *Soit* $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ *un polynome, où* $a_0, a_1, \ldots, a_n \in \mathbb{C}$.

Alors $p(x) = a_n(x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n)$, *où* $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{C}$.

De plus, si les a_i *sont tous réels, alors les* α_i *sont soit réels ou viennent en pairs conjugués.*

MAT 1741		**2. Vecteurs géométriques et algébriques**
Mike Newman	Université d'Ottawa	automne 2008

GÉOMETRIE

Un *scalaire* est une quantité ayant valeur seulement. Par exemple, $2, \sqrt{17}, \pi, 3 - i$ sont tous des scalaires. Par contre un *vecteur* possède longueur et direction. Par exemple, $\vec{u} =$ "3km au nord", et $\vec{v} =$ "2km 30 degrés au nord de l'est". Faites un dessin!

Un vecteur n'a pas de position fixe. On peut imaginer qu'il "commence" n'importe où. Ceci permet d'additionner des vecteurs de façon purement géométrique: la somme $\vec{u} + \vec{v}$ se trouve en déplacent $\vec{v}$ pour que son point de départ correspond au point terminal du $\vec{u}$. On peut aussi multiplier un vecteur par un scalaire (réel). Il s'agit de multiplier la longueur, et si le scalaire est négatif, inverser la direction. Donc $2\vec{v}$ est un vecteur 4km 30 degrés NE, et $-3\vec{u}$ est un vecteur 3km au sud.

COORDONNÉES

On présume en fait une sorte de "base". C'est-à-dire, on présume qu'on connait la direction "nord", "est", etc., ainsi que la distance "km". Ceci donne une autre point de vue. Si on pose le point de pépart de $\vec{v}$ à l'origine, on trouve que le point terminal se trouve au point $(\sqrt{3}, 1)$ (c'est notre ami Pythagore!). Donc, le vecteur $\vec{v}$ peut se décrire comme étant le vecteur allant de l'origine au point $(\sqrt{3}, 1)$. On écrit alors

$$\vec{v} = \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}$$

On peut donc faire l'arithmétique par coordonnées:

$$\vec{u} + \vec{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{3} \\ 2 \end{bmatrix} - 3\vec{v} = \begin{bmatrix} -3\sqrt{2} \\ -3 \end{bmatrix}$$

On dit que $\vec{u}$ et $\vec{v}$ sont des vecteurs en $\mathbb{R}^2$. Ceci signifie deux dimensions, avec coordonnées réelles. On peut aussi imaginer des vecteurs en $\mathbb{R}^3$: penser a l'espace qu'on habite. Par example, un vecteur 2km de long, en direction nord-ouest, faisant un angle de 30 degrés entre l'horizontal et le vertical. L'avantage d'un systéme de coodonnées et que ce vecteur ce décrit par une colonne de *trois* nombres réels. Si "est" correspond à l'axe des x,"nord" à l'axe des y, et "vertical" à l'axe des z, on a le point $(-\sqrt{3/2}, \sqrt{3/2}, 1)$ (c'est une exercice en trigonométrie!). On a le vecteur algébrique $\begin{bmatrix} -\sqrt{3/2} \\ \sqrt{3/2} \\ 1 \end{bmatrix}$.

La beauté c'est qu'on peut faire les calculs arithmétiques par coordonnées: on n'a pas besoin de faire des dessins en trois dimensions. Mais ce n'est pas le fin: pourquoi pas quatre? cinq? Ou n, pour n'importe quel entire $n > 0$?

On travail facilement avec des vecteurs algébriques dans n'importe quelle dimension. Mais la géometrie devient difficile pour $n > 3$! Par contre, notre intuition géométrique dans $\mathbb{R}^2$ et $\mathbb{R}^3$ peut nous aider pour $\mathbb{R}^n$ aussi.

Parfois la notation devient longue. Par exemple:

$$\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} \in \mathbb{R}^3 \qquad \begin{bmatrix} 4 \\ 6 \\ -4 \\ 0 \\ -8 \end{bmatrix} \in \mathbb{R}^5 \qquad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

Donc on voit parfois:

$$(4, 6, -4, 0, -8) \in \mathbb{R}^5$$

mais on la confond avec les coordonnées d'un point: je m'en servirait rarement. Aussi on a la notation du *transpose*

$$\begin{bmatrix} 4 & 6 & -4 & 0 & -8 \end{bmatrix}^T \in \mathbb{R}^5$$

Le "T" veut tout simplement dire "échanger les rangées et les colonnes" — si vous voulez, refleter le papier sur un axe diagonale!.

ARITHMÉTIQUE

Quelques règles. Notes qu'ils sont valides en $\mathbb{R}^2$, $\mathbb{R}^3$, $\ldots$, $\mathbb{R}^n$, et même pour $\mathbb{C}^n$ (on n'étudierait pas $\mathbb{C}^n$ dans ce cours: la géometrie est trop compliquée).

- $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
- $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$
- $\vec{u} + \vec{0} = \vec{u}$
- $\vec{u} + \vec{-u} = \vec{0}$ (le vecteur $\vec{0}$!)
- $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$ pour $a \in \mathbb{R}$
- $(a + b)\vec{u} = a\vec{u} + b\vec{u}$ pour $a, b \in \mathbb{R}$
- $a(b\vec{u}) = (ab)\vec{u}$ pour $a, b \in \mathbb{R}$
- $1\vec{u} = \vec{u}$ (le chiffre 1!)

Le vecteur $\vec{0}$ est un peu particulier. C'est le vecteur ayant tout coordonnées égal à zéro. C'est le seul vecteur n'ayant aucune direction!

Exercice: Démontrer ces formules pour $\mathbb{R}^2$. Mettre $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \ldots$

COMBINAISON LINÉAIRE

Étant donné des vecteurs $\vec{a}, \vec{b}, \vec{c}$, on dit que $\vec{x}$ est *combinaison linéaire* de $\vec{a}, \vec{b}, \vec{c}$ s'il existe des scalaires α, β, γ tel que

$$\vec{x} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c}$$

Comme petit exemple, on voit que $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ est combinaison linéaire de $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Les scalaires en question sont $1, 2, 3$, qu'on a trouvé en résolvant le système suivant:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Cést un exemple simple, on verra plus au DGD! Et dans la suite du cours aussi...

PRODUIT SCALAIRE

Le produit scalaire pour deux dimensions est $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = u_1v_1 + u_2v_2$ La formule se généralise facilement:

$$\boxed{\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + \cdots u_nv_n}$$

NB: Le produit scalaire donne une réponse *scalaire*. C'est un nom bien-choisi!

Ça sert à quoi? Premièrement, on calcul $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = u_1^2 + u_2^2$. Un dessin (Pythagore encore!) montre qu'en $\mathbb{R}^2$, $\vec{u} \cdot \vec{u}$ donne le carré de la longueur de $\vec{u}$. On dénote la longueur de $\vec{u}$ par $\|\vec{u}\|$. C'est peut-être moins connu que la formule de Pythagore fonctionne en toute dimension. On a donc que

$$\vec{u} \cdot \vec{u} = u_1^2 + u_2^2 + \cdots + u_n^2 = \|\vec{u}\|^2 .$$

NB: Il y a des similarités avec les complexes! On peut comprendre $\mathbb{C}$ commne étant une version spécialisée de $\mathbb{R}^2$ ...

On résume quelques propriétés.

- $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
- $\vec{u} \cdot \vec{0} = 0$ (un vecteur $\vec{0}$ et un scalaire 0!)
- $(a\vec{u}) \cdot \vec{v} = a(\vec{v} \cdot \vec{u}) = \vec{v} \cdot (a\vec{u})$ pour $a \in \mathbb{R}$
- $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$
- $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2$

Exercice: Vérifier ces formules pour $\mathbb{R}^2$. Mettre $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ ...

Soit $\vec{x}$ et $\vec{y}$ deux vecteurs. C'est souvent utile de "décomposer" $\vec{x}$ par rapport à $\vec{y}$. De façon précise, on cherche à écrire $\vec{x} = \vec{u} + \vec{v}$, où $\vec{u}$ est parallèle à $\vec{y}$ et $\vec{v}$ est perpendiculaire à $\vec{y}$. Voir le dessin en classe ...

En mettant le point de départ de $\vec{x}$ et $\vec{y}$ au même endroit, on voit que $\vec{u}$ et $\vec{v}$ existent, et sont uniques. On écrit typiquement $\text{proj}_{\vec{y}}(\vec{x})$ pour $\vec{u}$: c'est la partie de $\vec{x}$ dans la direction de $\vec{y}$. Par exemple, un dessin montre que si $\vec{x} = \begin{bmatrix} 3 & 7 \end{bmatrix}^T$ et $\vec{y} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$ alors $\text{proj}_{\vec{y}}(\vec{x}) = \begin{bmatrix} 3 & 0 \end{bmatrix}^T$.

En générale on a la formule suivante (qui peut se démontrer à l'aide du loi de cosinus, par exemple)

$$\boxed{\text{proj}_{\vec{y}}(\vec{x}) = \frac{\vec{x} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} \vec{y} = \frac{\vec{x} \cdot \vec{y}}{\|y\|^2} \vec{y}}$$

NB: La fraction est un *chiffre*, qui multiplie le vecteur $\vec{y}$ (multiplication scalaire!).

Exercice: Pour l'exemple ci-haut, calculer $\text{proj}_{\vec{y}}(\vec{x})$ afin de comparer avec la réponse géométrique. Calculer aussi $\text{proj}_{\vec{x}}(\vec{y})$.

On a comme but une décomposition $\vec{x} = \vec{u} + \vec{v}$ où $\vec{u}$ est parallèle à $\vec{y}$ et $\vec{v}$ est perpendiculaire à $\vec{y}$. On a $\vec{u} = \text{proj}_{\vec{y}}(\vec{x})$, mais comment trouver $\vec{v}$?

On note que $\vec{x} = \text{proj}_{\vec{y}}(\vec{x}) + (\vec{x} - \text{proj}_{\vec{y}}(\vec{x}))$ (n'est-ce pas?). Pourquoi pas prendre $\vec{v} = \vec{x} - \text{proj}_{\vec{y}}(\vec{x})$? Comment savoir si c'est perpendiculaire à $\vec{y}$? On trouvera la réponse bientôt...

Revenons au dessin de $\vec{x}$ et $\vec{y}$. On voit qu'il y a un angle θ entre ces deux vecteurs. Un triangle rectangle nous indique que

$$\begin{aligned}
\cos(\theta) &= \frac{\left\|\text{proj}_{\vec{y}}(\vec{x})\right\|}{\|\vec{x}\|} \\
&= \frac{\sqrt{\left(\frac{\vec{x}\cdot\vec{y}}{\vec{y}\cdot\vec{y}}\vec{y}\right) \cdot \left(\frac{\vec{x}\cdot\vec{y}}{\vec{y}\cdot\vec{y}}\vec{y}\right)}}{\sqrt{\vec{x} \cdot \vec{x}}} \\
&= \frac{\vec{x} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} \sqrt{\frac{\vec{y} \cdot \vec{y}}{\vec{x} \cdot \vec{x}}} \\
&= \frac{\vec{x} \cdot \vec{y}}{\sqrt{\vec{y} \cdot \vec{y}}\sqrt{\vec{x} \cdot \vec{x}}} = \frac{\vec{x} \cdot \vec{y}}{\|\vec{y}\| \, \|\vec{x}\|}
\end{aligned}$$

On trouve donc que

$$\boxed{\vec{x} \cdot \vec{y} = \|\vec{x}\| \, \|\vec{y}\| \cos \theta}$$

C'est important pour deux raisons: on a une relation entre la géometrie (l'angle) et l'algèbre (prduit scalaire). Une application directe c'est un test de perpendicularité.

Deux vecteurs sont perpendiculaires si et seulement si leur produit scalaire est zéro.

$$\boxed{\vec{x} \perp \vec{y} \iff \vec{x} \cdot \vec{y} = 0}$$

Si deux vecteurs sont perpendiculaires, on dit aussi qu'ils sont *orthogonaux*. On verra plus tard que d'autres choses que des vecteurs peuvent être orthogonaux...

Rappel: on se demandait si $\vec{x} - \mathrm{proj}_{\vec{y}}(\vec{x})$ et $\vec{y}$ sont ortohgonales. La réponse est "oui", car:

$$\vec{y} \cdot \left(\vec{x} - \frac{\vec{x} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} \vec{y} \right) = \vec{y} \cdot \vec{x} - \vec{y} \cdot \left(\frac{\vec{x} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} \vec{y} \right) = \vec{y} \cdot \vec{x} - \frac{\vec{x} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} \vec{y} \cdot \vec{y} = 0$$

De façon générale soit $\vec{x}$ et $\vec{y}$ deux vecteurs. Alors on peut écrire

$$\boxed{\vec{x} = \big(\mathrm{proj}_{\vec{y}}(\vec{x})\big) + \big(\vec{x} - \mathrm{proj}_{\vec{y}}(\vec{x})\big)}$$

La première parenthèse est un vecteur parallèle à $\vec{y}$; la deuxième, orthogonal à $\vec{y}$.

NB: On vient de déterminer une *base orthogonale*, qui sera un concept très important pour les espaces vectoriels. En fait, on a vu un le début de la *méthode de Gramm-Schmidt*... On s'en reparle!

DROITES ET PLANS

On connait ce que c'est une droite. Du point de vue vectoriel, il suffit de donner un point départ et une direction.

$$L = \left\{ \vec{v_0} + t\vec{d} \mid t \in \mathbb{R} \right\}$$

On la comprend comme suit: $\vec{v_0}$ est le vecteur correspondant au point de départ et $\vec{d}$ est un vecteur dans la direction de la droite.[1] On dit que $\vec{d}$ est *vecteur directeur*. Le paramètre t peut varier, afin de donner tous les points sur la droite. On dit que c'est une *équation paramétrique*.

Pour une seule droite on peut choisir des différents points de départ, ainsi que des différents vecteurs (pourvu qu'ils ont tous la même direction). Donc l'équation n'est pas unique:

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 2 \\ 3 \end{bmatrix} \mid t \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 3 \\ 4 \end{bmatrix} + s \begin{bmatrix} 2 \\ 3 \end{bmatrix} \mid t \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 7 \\ 10 \end{bmatrix} + r \begin{bmatrix} -4 \\ -6 \end{bmatrix} \mid t \in \mathbb{R} \right\}$$

Exercice: Quelle est la relation entre t, r et s?

Si un point (x, y) est sur la droite L, on a

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

pour une valeur de t, ce qui équivaut au système suivant:

$$\begin{cases} x = 1 + 2t \\ y = 1 + 3t \end{cases}$$

On peut écrire une équation *vectoriel* ou *paramétrique* en n'importe quelle dimension; l'exemple ci-haut étant pour $\mathbb{R}^2$.

Pour les plans, on traite un cas spécial: $\mathbb{R}^3$. Pour décrire un plan en $\mathbb{R}^3$, on donne un point de départ et un vecteur *normal*. C'est-à-dire, un vecteur orthogonal au

[1] On pense souvent en terms de points (x, y) sur L; par contre l'équation décrit des vecteurs $\begin{bmatrix} x \\ y \end{bmatrix}$. C'est l'équivalence entre un point et un vecteur en position standard. Les vecteurs sont plus utiles pour les calculs, même si c'est les points qui nous intéressent.

plan. Si le vecteur normal est $\begin{bmatrix} n_1 & n_2 & n_3 \end{bmatrix}^T$ et le plan passe par (p_1, p_2, p_3), alors le point (x, y, z) se trouve sur le plan si et seulement si on a

$$\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \right) \cdot \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = 0$$

C'est *l'équation normale* ou parfois *l'équation scalaire* d'un plan. On verra l'explication géométrique en classe... mais n'y atttendez pas: faites-vous un graphique!

Prenons un exemple. Le plan passant par $(-2, -1, 1)$ ayant comme normale $\begin{bmatrix} 2 & -2 & 7 \end{bmatrix}^T$ a l'équation

$$\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right) \cdot \begin{bmatrix} 2 \\ -2 \\ 7 \end{bmatrix} = 0.$$

En simplifiant (régles d'arithmétique du produit scalaire), on obtient:

$$2x - 2y + 7z = 5.$$

Les coefficients sont exactement les coordonnées du vecteur normal. Ce n'est pas par hasard.

Exercice: : Que signifie la constante, c'est-à-dire, le 5 dans l'exemple précédent?

GÉOMETRIE EN $\mathbb{R}^n$

Un petit coup-d'œuil vers la direction futur du cours...

Qu'arrive-t-il si on adapte l'équation normale d'un plan au cas de $\mathbb{R}^2$? Par exemple, prennons un point de départ $(1, 1)$ et un vecteur normal $\begin{bmatrix} 3 & -2 \end{bmatrix}^T$. Ceci donne

$$\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \cdot \begin{bmatrix} 3 \\ -2 \end{bmatrix} = 0$$

ou plutot $3x - 2y = 1$. C'est l'équation d'une droite!

Exercice: Quelle est la relation entre cette droite est la première exemple ci-haut? (indice: Quelle est l'angle entre le vecteur normal ici et le vecteur directeur ci-haut?)

Une équation paramatrique pour une droite donne un point de départ et une direction permise. C'est valide en tout $\mathbb{R}^n$.

Une équation normale donne un point de départ et une direction défendue. En $\mathbb{R}^3$, c'est un plan; en $\mathbb{R}^2$, c'est une droite; en $\mathbb{R}^n$...?

Comme d'habitude, l'approche algébrique permet de généraliser. On peut imaginer *plusieurs* directions permises ou défendues. On verra ceci plus tard dans le cours.

Question: On a vu une équation paramétrique et une éqation normale pour une même droite en $\mathbb{R}^2$ (exercice ci-haut!). On avait **1** direction permise, **1** direction défendue, et **2** dimensions en $\mathbb{R}^2$, et $1 + 1 = 2$. Hasard?

Question: Est-ce vrai que l'exemple du plan ci-haut équivaut à

$$\left\{ \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 7 \\ 2 \end{bmatrix} \mid t, s \in \mathbb{R} \right\}$$

Ce sont de questions à revisiter plus tard.

APPLICATIONS

On traite deux sortes de questions: l'intersection de droites et l'intersection de plans.

Problème: Trouver l'intersection des deux droites

$$L_1 = \left\{ \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\} \qquad L_1 = \left\{ \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} + t \begin{bmatrix} -3 \\ -1 \\ 0 \end{bmatrix} \right\}$$

On cherche une valuer de t dans la première, ainsi qu'une valeur de t dans la deuxième qui donne le même point. Mais attention: il n'y a aucune raison de penser que ce seront les *mêmes* valeurs de t. Il y a vraiament *deux* paramètres... donc on leur donne deux noms différents, par exemple s et t. On cherche alors des valeurs t et s tel que

$$\begin{aligned} 2 + 0t &= -1 - 3s \\ 1 + t &= 1 - s \\ 3 + 2t &= 5 + 0s \end{aligned}$$

La première équation donne $s = -1$, et la dernière donne $t = 1$. **Attention!** On n'a pas encore fini. On cherche une solution au *système*. Il faut vérifier qu'on a une solution pour toutes les *trois* équations. On vérifie que la deuxième est satisfaite, car $1 + (1) = 1 - (-1)$. Donc $t = 1$, $s = -1$ est la solution unique.[2] On subsititue $t = 1$ dans la première droite pour donner le point $(2, 2, 5)$; également on substitue $s = -1$ pour t dans la deuxième pour donner le point $(2, 2, 5)$. Bien sûr, c'est le même point...

Étant donné que deux droites intersectent, ils définissent un plan. Comment trouver ce plan?

On connait déja un point sur ce plan: $(2, 2, 5)$. Mais comment trouver un vecteur normal? On laissera cette question pour quelques instants, mais on donnera la réponse toute suite: $\begin{bmatrix} 2 & -6 & 3 \end{bmatrix}^T$.

[2]En résolvant deux équations, on a montré que *si* il existe une solution, *alors* la solution doit être $t = 1$, $s = -1$. En vérifiant l'autre équation, on a démontrer que oui, c'est une solution, qui est donc unique.

Exercice: Vous ne savez peut-être pas *comment* j'ai calculé ce vecteur, mais vous êtes capable de vérifier que c'est un vecteur orthogonal au deux droites. Allez-y! (indice: produit scalaire)

On a donc l'équation de notre plan:

$$\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}\right) \cdot \begin{bmatrix} 2 \\ -6 \\ 3 \end{bmatrix} = 0$$

ou plutôt $2x - 6y + 3z = 7$.

Si deux droites intersectent en un point, ils sont contenues dans un plan unique. S'ils n'intersectent pas, ils ne sont contenues dans aucun plan. Il y a aussi le cas où elles sont parallèles: là aussi, il y a un plan unique. Ce sont les seules possibilités en $\mathbb{R}^n$.

Problème: Trouver l'intersection des deux plans $x + y + z = 2$ et $x + 2y + 3z = 0$.

Du point de vue géométrique, c'est l'intersection de deux plans. Algébriquement, c'est un système d'équations. En éliminant x on obtient $y - 2z = -2$. En éliminant y on obtient $-x + z = -4$. On ne peut rien éliminer de plus, alors on a l'idée que la solution est

$$\begin{aligned} y &= -2 + 2z \\ x &= 4 + z \end{aligned}$$

La "solution" est en autre système d'équations? Oui: la variable z fonctionne comme *paramètre*: elle peut prendre n'importe qulle valeur réelle. Mettons donc $z = t$. On a alors une équation paramétrique

$$\begin{aligned} x &= 4 + t \\ y &= -2 + 2t \\ z &= t \end{aligned}$$

ou plutôt

$$\left\{ \begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \mid t \in \mathbb{R} \right\}.$$

La solution est une droite.

Exercice: Qu'arrive-t-il si on élimine x et z? Si on élimine y et z? On espère que ça revient au même?

On cherchait tantôt un vecteur orthogonal à deux vecteurs donnés. Comment là calculer? En $\mathbb{R}^3$ il y a une manière simple[3]: le produit vectoriel.

Définition 1.

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_2v_3 - v_2u_3 \\ -u_1v_3 + v_1u_3 \\ u_1v_2 - v_1u_2 \end{bmatrix}$$

On vérifie facilement que $\vec{u} \times \vec{v}$ est orthogonal à $\vec{u}$:

$$\begin{aligned} \begin{bmatrix} u_2v_3 - v_2u_3 \\ -u_1v_3 + v_1u_3 \\ u_1v_2 - v_1u_2 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} &= u_1(u_2v_3 - v_2u_3) + u_2(-u_1v_3 + v_1u_3) + u_3(u_1v_2 - v_1u_2) \\ &= u_1u_2v_3 - u_1v_2u_3 - u_1u_2v_3 + v_1u_2u_3 + u_1v_2u_3 - v_1u_2u_3 \\ &= 0 \end{aligned}$$

La vérification pour $\vec{v}$ est similaire. Donc $\vec{u} \times \vec{v}$ est un vecteur orthogonal à $\vec{u}$ et $\vec{v}$.

Il y a manière de comprendre cette formule en termes de *déterminants* (on verra les déterminants en détail plus tard). On commence en nommant trois vecteurs spéciaux:

$$\hat{\imath} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \hat{\jmath} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \hat{\mathrm{k}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Entre autre, ceci permet d'écrire $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x\hat{\imath} + y\hat{\jmath} + z\hat{\mathrm{k}}$.

Pour le produit vectoriel, on a la formule en termes de déterminants:

$$\begin{aligned} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} &= \begin{vmatrix} \hat{\imath} & u_1 & v_1 \\ \hat{\jmath} & u_2 & v_2 \\ \hat{\mathrm{k}} & u_3 & v_3 \end{vmatrix} \\ &= \hat{\imath}\,(u_2v_3 - v_2u_3) - \hat{\jmath}\,(u_1v_3 - v_1u_3) + \hat{\mathrm{k}}\,(u_1v_2 - v_1u_2) \\ &= \begin{bmatrix} u_2v_3 - v_2u_3 \\ -u_1v_3 + v_1u_3 \\ u_1v_2 - v_1u_2 \end{bmatrix} \end{aligned}$$

Si vous ne connaissez pas les déterminants, utilisez la formule directe.

[3]En R^n, il y a aussi une manière "simple". Le problème, c'est qu'il n'y a pas une réponse unique!

Quels sont les propriétés du produit vectoriel?

On peut facilement calculer, par exemple, que $\hat{\imath} \times \hat{\jmath} = \hat{k}$, mais $\hat{\jmath} \times \hat{\imath} = -\hat{k}$. En générale, on trouve que

- $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$

On a aussi les règles suivants:

- $\vec{u} \times \vec{0} = \vec{0}$
- $\vec{u} \times \vec{u} = \vec{0}$
- $(a\vec{u}) \times \vec{v} = a(\vec{u} \times \vec{v}) = \vec{u} \times (a\vec{v})$
- $(\vec{u} \times \vec{v}) \times \vec{w} = \vec{u} \times (\vec{v} \times \vec{w})$
- $(\vec{u} + \vec{v}) \times \vec{w} = (\vec{u} \times \vec{w}) + (\vec{v} \times \vec{w})$

Exercice: Vous pouvez vérifiez ces règles en $\mathbb{R}^3$: mettre $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$, et faire les calculs. Par exemple

$$\vec{u} \times \vec{0} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} u_2 \cdot 0 - 0 \cdot u_3 \\ u_1 \cdot 0 - 0 \cdot u_3 \\ u_1 \cdot 0 - 0 \cdot u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}$$

AIRE ET VOLUME

(cette section est optionnelle)

On découvre une relation entre les produits scalaires et vectoriels en $\mathbb{R}^3$. Premièrement:

$$\begin{aligned}
\|\vec{u} \times \vec{v}\|^2 &= \left\| \begin{bmatrix} u_2v_3 - v_2u_3 \\ -u_1v_3 + v_1u_3 \\ u_1v_2 - v_1u_2 \end{bmatrix} \right\|^2 \\
&= (u_2v_3 - v_2u_3)^2 + (-u_1v_3 + v_1u_3)^2 + (u_1v_2 - v_1u_2)^2 \\
&= u_2^2v_3^2 - 2u_2v_3v_2u_3 + v_2^2u_3^2 + u_1^2v_3^2 - 2u_1v_3v_1u_3 + v_1^2u_3^2 + u_1^2v_2^2 - 2u_1v_2v_1u_2 + v_1^2u_2^2 \\
&= u_1^2v_2^2 + u_1^2v_3^2 + u_2^2v_1^2 + u_2^2v_3^2 + u_3^2v_1^2 + u_3^2v_2^2 - 2(u_1u_2v_1v_2 + u_1u_3v_1v_3 + u_2u_3v_2v_3) \\
&= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2
\end{aligned}$$

Ouf! La dernière égalité se voit peut-être plus facilement à l'envers: faites l'expansion de la dernière ligne.

Par contre, on connait déja que:

$$\begin{aligned}
\|\vec{u}\| &= u_1^2 + u_2^2 + u_3^2 \\
\|\vec{v}\| &= v_1^2 + v_2^2 + v_3^2 \\
\vec{u} \cdot \vec{v} &= u_1v_1 + u_2v_2 + u_3v_3
\end{aligned}$$

On a découvert que $\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\| \, \|\vec{v}\| - (\vec{u} \cdot \vec{v})^2$, ou le

Théorème 2.

$$\|\vec{u} \times \vec{v}\|^2 + (\vec{u} \cdot \vec{v})^2 = \|\vec{u}\| \, \|\vec{v}\|$$

C'est un peu comme le théorème de Pythagore. En fait, sachant que $\vec{u} \cdot \vec{v} = \|\vec{u}\| \, \|\vec{v}\| \cos\theta$ on obtiens alors que

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin\theta$$

Ceci donne des conséquences géomteriques. Deux vecteurs $\vec{u}$ et $\vec{v}$ définissent un paralléllogramme, si on met leurs points de départ ensemble. Si on prend $\|\vec{v}\|$ comme la base, alors la hauteur sera $\|\vec{u}\| \sin\theta$, donnant que l'aire est exactement $\|\vec{u} \times \vec{v}\|$.

On peut continuer. Trois vecteurs $\vec{u}$, $\vec{v}$ et $\vec{w}$ définissent un *parallélépidède*. C'est un objet en trois dimensions, un peu comme une brique: les faces opposées sont parallèles, mais les angles ne sont pas nécessairement de 90°. Le volume d'un parallélépidède est la surface de la base multipliée par la hauteur. Prennons la base comme étant la face définit par $\vec{u}$ et $\vec{v}$. Alors on connait la surface de la base: c'est $\|\vec{u} \times \vec{v}\|$. La hauteur est la longeur de la projection du vecteur $\vec{w}$ sur la normale, ou $\left\|\mathrm{proj}_{\vec{u}\times\vec{v}}(\vec{w})\right\|$. Donc le volume est

$$\begin{aligned} V &= \left\|\mathrm{proj}_{\vec{u}\times\vec{v}}(\vec{w})\right\| \, \|\vec{u} \times \vec{v}\| \\ &= \left\| \frac{\vec{u} \cdot (\vec{u} \times \vec{v})}{\|\vec{u} \times \vec{v}\|^2} (\vec{w} \times \vec{v}) \right\| \, \|\vec{u} \times \vec{v}\| \\ &= \left| \frac{\vec{u} \cdot (\vec{u} \times \vec{v})}{\|\vec{u} \times \vec{v}\|^2} \right| \, \|\vec{w} \times \vec{v}\| \, \|\vec{u} \times \vec{v}\| \\ &= |\vec{w} \cdot (\vec{u} \times \vec{v})| \end{aligned}$$

Il n'y a aucune raison de prendre $\vec{u}$ et $\vec{v}$ commen base: on aurait pu prendre les trois vecteurs en n'importe quel ordre. Donc

$$|\vec{w} \cdot (\vec{u} \times \vec{v})| = |\vec{u} \cdot (\vec{v} \times \vec{w})| = |\vec{v} \cdot (\vec{u} \times \vec{w})| = \cdots$$

Finalement, on note que

$$\vec{u} \cdot (\vec{v} \times \vec{w}) = \left| \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \right|$$

et donc le volume est la valeur absolue du déterminant.

ESPACES VECTORIELS

On a déja vu les vecteurs dans un context géométrique, ainsi que algébrique: $\mathbb{R}^2$, $\mathbb{R}^3$, et en généralisant, $\mathbb{R}^n$. On a vu des propriétés (tel que $\vec{u}+\vec{v}=\vec{v}+\vec{u}$) ainsi que des méthodes (tel que projection). Voyons maintenant quelques examples un peu différents. . .

EXAMPLE: ÉQUATIONS

Considérons les équations linéaires suivants:

$$\begin{aligned} E_1 :&\quad x+2y-z=3 \\ E_2 :&\quad 2x-y+5z=1 \\ E_3 :&\quad x+y+z=0 \end{aligned}$$

On pourra demander des questions tel "résoudre le système", mais on cherche à comprendre ces objets d'une autre manière, en fait, on veut les considérer comme objets en soi.

On peut *additionner* des équations. Par exemple E_1+E_2 donne l'équation $3x+y+4z=4$. On peut les multiplier par un scalaire: $5E_3$ donne $-x+5y+5z=0$. On peut aussi prendre la négation: $-E_2$ donne $-2x+y-5z=-1$. On voit que $E_2+(-E_2)$ donne. . . quoi exactement? Si on suit la logique, on obtient $0x+0y+0z=0$, c'est-a-dire, l'équation $0=0$.

Les équations se comportent un peu comme les "vecteurs".

Comparons avec les "règles d'arithmétique" des vecteurs vu antérieurement. On dénote par $\mathcal{E}_k$ l'ensemble de toutes les équations linéaires en k variables. Soit E_1, E_2 et E_3 des éléments de $\mathcal{E}_3$, un peu comme les exemples ci-haut.

On voit directement que:

- $E_1+E_2=E_2+E_1$
- $E_1+(E_2+E_3)=(E_1+E_2)+E_3$

Si on définit l'équation $0=0$ comme Z, on voit que:

- $E_1+Z=E_1$
- $E_1+(-E_1)=Z$

La multiplication scalaire est distributative et associative:

- $a(E_1+E_2)=aE_1+aE_2$

- $(a+b)E_1 = aE_1 + bE_1$
- $a(bE_1) = (ab)E_1$

Finalement, on voit que le scalaire 1 se distingue:

- $1E_1 = E_1$

Exercice: Vérifier chaque énoncé ci-haut pour les équations données au début de la section, ainsi que les valuers $a = 3, b = -5$.

NB: On n'a nul part fait mention de ces équations comme étants "vraies" ou "fausses", si elles sont satisfaites ou non. Ce sont des objets.

EXAMPLE: FONCTIONS

Considérons l'ensemble de tous les fonctions réelles:

$$\mathcal{F} = \{f : \mathbb{R} \to \mathbb{R}\}$$

Soit f, g deux fonctions en $\mathcal{F}$. On peu définir la fonction $f + g$ en disant que l'évaluation de $f + g$ à la valeur x est la somme de l'évaluation de f à x et de l'évaluation de g à x. En formule, $(f + g)(x) = f(x) + g(x)$. On pourrait aussi définir une multiplication scalaire. Si $a \in \mathbb{R}$, alors af est la fonction pour laquelle l'évaulation à la valeur x est a fois l'évaluation de la fonction f à la valeur x. En formule, $(af)(x) = af(x)$

Un exemple est peut-être requis! Si $f(x) = x^2$ et $g(x) = \sin(x)$, alors $(f + g)(x) = x^2 + \sin(x)$ et $(3f)(x) = 3x^2$. Vous vous demandez peut-être en ce moment... "M'enfin?!"

Vous auriez raison: on n'a rien inventé de nouveau. Mais on a découvert un autre point de vue. On peut traiter les fonctions comme des objets, et ce étant fait, elles se comportent un peu comme des vecteurs. "Comme des vecteurs"? Voyons, par exemple, la commutativité de l'addition et une distributativité de la multiplication scalaire.

$$f + g = x^2 + \sin(x) = \sin(x) + x^2 = g + f$$
$$(2 + 3)f = 5x^2 = 2x^2 + 3x^2 = 2f + 3f$$

L'arithmétique de fonctions est un peu comme l'arithmétique des vecteurs géométriques. Comparons avec les règles d'arithmétique des vecteurs. Soit f_1, f_2 et f_3 des fonctions dans $\mathcal{F}$.

On a la commutativité et l'associativité de l'addition:

- $f_1 + f_2 = f_2 + f_1$
- $f_1 + (f_2 + f_3) = (f_1 + f_2) + f_3$

Si on définit la fonction $Z(x) = 0$, on a:

- $f_1 + Z = f_1$

- $f_1 + (-f_1) = Z$

La multiplication scalaire est distributative et associative:

- $a(f_1 + f_2) = af_1 + af_2$
- $(a + b)f_1 = af_1 + bf_1$
- $a(bf_1) = (ab)f_1$

Finalement, le scalaire 1 se distingue:

- $1f_1 = f_1$

Exercice: Vérifier chaque énoncé ci-haut pour les fonctions $f_1(x) = x^2$, $f_2(x) = \sin(x) - x^2$, et $f_3(x) = 2x + x^2$, ainsi que $a = 2, b = -2$.

NB: On n'a nul part fait mention de la valeur numérique de ces fonctions à une valeur particulière de x. L'évaluation de sin(123) ne nous intéresse pas du tout. Ce sont des objets.

Il reste une autre propriété des vecteurs qu'on n'a pas encore mentionné. La somme de deux vecteurs *est* un vecteur. C'est pareil pour les fonctions: $f + g$ est aussi une fonction: pour chaque valeur de x, $f + g$ possède une valeur unique: c'est exactement ce que veut dire la notation "$(f + g)(x) = f(x) + g(x)$".

Question: Définissons les ensembles de fonctions $\mathcal{G} = \{f : \mathbb{R} \to \mathbb{R} \mid f(0) = 0\}$ et $\mathcal{H} = \{f : \mathbb{R} \to \mathbb{R} \mid f(0) = 1\}$. Est-ce que ces ensembles se comportent comme des vecteurs?

EXEMPLE: SOUS-ENSEMBLES DE $\mathbb{R}^n$

Considérons l'ensemble des points dans $\mathbb{R}^2$ définit comme

$$U = \{(x, 2x) \mid x \in \mathbb{R}\}.$$

On connait déja comment additionner des éléments de U, ainsi de comment les multiplier par des scalaires: ce sont des propriétés hérités de $\mathbb{R}^2$.

Voici trois points qui font parti de l'ensemble U: $(3, 6)$, $(-1, -2)$ et $(5, 10)$. On a donc $(3, 6) + (-1, -2) = (2, 4)$ et $5(-1, -2) = (-5, -10)$. On peu aussi considérer les vecteurs (en position standard!): $\begin{bmatrix} 3 \\ 6 \end{bmatrix} + \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, etc.

On voit que les règles d'arithmétiques de vecteurs s'appliquent directement à U, puisque les éléments de U sont déja des vecteurs!

On vérifie de plus que si $\vec{u} = \begin{bmatrix} u_1 \\ 2u_1 \end{bmatrix}$ et $\vec{v} = \begin{bmatrix} v_1 \\ 2v_2 \end{bmatrix}$ sont deux éléments de U, on a $\vec{u} + \vec{v} \in U$, ainsi que $a\vec{u} \in U$ pour n'importe quelle $a \in \mathbb{R}$.

Changeons un peu l'exemple:

$$W = \{(x, 2x + 1) \mid x \in \mathbb{R}\}.$$

Les régles d'arithmétique sont encore satisfaits. Par contre, quelque chose de bizarre se produit. Prennons $\begin{bmatrix}1\\3\end{bmatrix}$ et $\begin{bmatrix}-7\\-13\end{bmatrix}$ dans W. Leur somme $\begin{bmatrix}1\\3\end{bmatrix}+\begin{bmatrix}-7\\-13\end{bmatrix}=\begin{bmatrix}-6\\-10\end{bmatrix}$ n'est pas un élement de W. De plus $2\begin{bmatrix}1\\3\end{bmatrix}=\begin{bmatrix}2\\6\end{bmatrix}$ ne l'est pas non plus.

La différence essentiel est que l'arithmétique dans U reste dans U, mais l'arithmétique de W peut en sortir: U est "fermé", mais W ne l'est pas.

EXEMPLE: MATRICES

On définit une *matrice réelle de taille* $m\times n$ comme étant un objet ayant m rangées, n colonnes, et pour chaque position, une valeur réelle. On dénote par $\mathbb{M}_{m,n}(\mathbb{R})$ l'ensemble de toutes les matrices réelles à m rangées et n colonnes ayant des valeurs réelles à chaqie position. Puisqu'on traitera (presque) toujours de matrices réelles, on écrit alors $\mathbb{M}_{m,n}$. Par exemple,

$$\begin{bmatrix}1&0\\0&1\end{bmatrix}\in\mathbb{M}_{2,2}\qquad\begin{bmatrix}1&2&3\\0&-1&\pi\end{bmatrix}\in\mathbb{M}_{2,3}\qquad\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix}\in\mathbb{M}_{3,3}$$

On peut additionner deux matrices, si elles ont la même taille. La somme est calculée par position. Par exemple,

$$\begin{bmatrix}0&2&4\\1&1&0\end{bmatrix}+\begin{bmatrix}1&2&3\\0&-1&\pi\end{bmatrix}=\begin{bmatrix}0+1&2+2&4+3\\1+0&1+(-1)&1+\pi\end{bmatrix}=\begin{bmatrix}1&4&7\\1&0&1+\pi\end{bmatrix}$$

Si les matrices n'ont pas la même taille, on ne peut pas les additionner. On voit directement que $A+B=B+A$ pour toutes matrices A,B de la même taille.

On peu aussi définir une multiplication scalaire, par position. Par exemple:

$$3\begin{bmatrix}1&2&0\\0&-5&3\end{bmatrix}=\begin{bmatrix}3\times 1&3\times 2&3\times 0\\3\times 0&3\times(-5)&3\times 3\end{bmatrix}=\begin{bmatrix}3&6&0\\0&-15&9\end{bmatrix}$$

Vous dites peut-être qu'un vecteur, c'est nul autre qu'une matrice n'ayant une seule colonne! Ce n'est pas tout-à-fait correct... Par contre, l'arithmétique des matrices est très similiare à l'arithmétique des vecteurs. (On dirait presque qu'il y a un thème à ce chapitre!)

Prenons trois matrices A,B,C en $\mathbb{M}_{2,3}$ (le choix de 2 et 3 n'est pas vraiment important...).

En regardant attentivement à notre exemple ci-haut, on voit que l'addition est commutative. En fait, elle est commutative exactement parce-que l'addition des nombres réelles est commutative. L'associativité est hérité des réelles de façon similaire:

- $A+B=B+A$
- $A+(B+C)=(A+B)+C$

Si on définit la matrice $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ comme Z, on voit que:

- $A + Z = A$
- $A + (-A) = Z$

La multiplication scalaire est distributative et associative:

- $a(A + B) = aA + aB$
- $(a + b)A = aA + bA$
- $a(bA) = (ab)A$

Le scalaire 1 se distingue:

- $1A = A$

Exercice: Comme d'habitude, on laisse la vérification de ces règles comme exercice

POLYNÔMES

Définissons $\mathbb{P}_k$ comme l'ensemble de tout polynôme de degré au plus k. Par exemple, $p_1 = 3x + 4$ et $p_2 = 5x^2 - 2x$ sont dans $\mathbb{P}_2$.

On peut définir une addition de polynômes, ainsi qu'une multiplication scalaire:

$$p_1 + p_2 = (0 + 5)x^2 + (3 - 2)x + (4 + 0) = 5x^2 + x + 4$$
$$17p_1 = 17(5)x^2 + 17(-2)x + 5(0) = 85x^2 - 34x$$

Question: On se demande si les polynômes se comportent un peu comme les vecteurs. Est-ce que les "mêmes" règles d'arithmétique s'appliquent? La question n'est pas complètement banale: pourquoi a-t-on dit "degré *au plus k*" et non "degré *égal à k*".

AXIOMES. . .

On voudrait un concept pour réunir tous ces exemples. C'est l'idée d'une *espace vectoriel*.

Définition 1. *Un espace vectoriel réelle est en ensemble d'objets V ainsi que l'ensemble de scalaires*[1] $\mathbb{R}$*, muni d'une addition et d'une multiplication scalaire, tel que les axiomes suivants sont tous satisfaits pour* tout $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ *et pour* tout $a, b \in \mathbb{R}$.

On a deux axiomes de fermeture:

1) $\mathbf{x} + \mathbf{y} \in V$
2) $a\mathbf{x} \in V$

On a les règles d'arithmétique pour l'addition:

3) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
4) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$

Ainsi que pour la multiplication scalaire:

5) $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$
6) $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$
7) $a(b\mathbf{x}) = (ab)\mathbf{x}$

On a un zéro et des inverses:

8) Il existe un vecteur $\mathbf{0} \in V$ *tel que* $\mathbf{x} + \mathbf{0} = \mathbf{x}$
9) Pour chaque $\mathbf{x} \in V$ *il existe un* $-\mathbf{x} \in V$ *tel que* $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$

On a dernièrement un scalaire spécial aussi, le chiffre 1

10) $1\mathbf{x} = \mathbf{x}$

On écrit souvent les vecteurs en gras, comme $\mathbf{x}$. C'est pour faire la distinction entre un vecteur dans un espace vectoriel quelconque, est un vecteur géométrique en $\mathbb{R}^n$.

NB: Attention au $\mathbf{0}$. Le vecteur $\mathbf{0}$ et le chiffre 0 ne sont pas la même chose du tout!

Afin de démontrer que un ensemble V est un espace vectoriel, il faut montrer que chaque axiome est satisfaite. Si un seul axiome n'est pas satisfait pour un seul vecteur, alors ce n'est pas un espace vectoriel.

À titre d'exemple, voici une démonstration de l'axiome 4) pour $\mathbb{M}_{2,2}$. On prend trois éléments arbitraires dans $\mathbb{M}_{2,2}$, et on montre que les deux côtés de l'égalité sont égales:

[1]Pour nous les scalaires seront les nombres réelles $\mathbb{R}$. En générale on peut remplacer $\mathbb{R}$ par n'importe quel *corps*, par exemple $\mathbb{Q}$ ou $\mathbb{C}$. Il y a aussi des corps finis, par exemple $\mathbb{Z}_p$, p un nombre premier (et d'autres!): qui sont très importants pour le codage (transmission d'information de manière robuste) et la cryptographie (transmission d'information de manière secrète). On n'a pas le temps d'en parler dans ce cours!

$$\begin{aligned}
&\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \left(\begin{bmatrix} e & f \\ g & h \end{bmatrix} + \begin{bmatrix} i & j \\ k & l \end{bmatrix} \right) \\
&\qquad = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e+i & f+j \\ g+k & h+l \end{bmatrix} && \text{addition en } \mathbb{M}_{2,3} \\
&\qquad = \begin{bmatrix} a+(e+i) & b+(f+j) \\ c+(g+k) & d+(h+l) \end{bmatrix} && \text{addition en } \mathbb{M}_{2,3} \\
&\qquad = \begin{bmatrix} (a+e)+i & (b+f)+j \\ (c+g)+k & (d+h)+l \end{bmatrix} && \text{addition en } \mathbb{R} \\
&\qquad = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix} + \begin{bmatrix} i & j \\ k & l \end{bmatrix} && \text{addition en } \mathbb{M}_{2,3} \\
&\qquad = \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right) + \begin{bmatrix} i & j \\ k & l \end{bmatrix} && \text{addition en } \mathbb{M}_{2,3}
\end{aligned}$$

C'est un peu long! Mais on voit qu'il n'y a pas d'idées compliquées, il s'agit de suivre les définitions. Notons que l'axiome 8) est un peu différent, car on doit deviner *qu'est-ce que c'est* le vecteur $\mathbf{0}$. Ce n'est pas nécessairement "juste des zéros" (c'est une indice pour votre prochain devoir!), mais souvent c'est le cas. Pour l'exemple de $\mathbb{M}_{2,3}$ on pense que peut-être $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Vérifions, avec $\mathbf{x} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ arbitraire:

$$\mathbf{x} + \mathbf{0} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a+0 & b+0 \\ c+0 & d+0 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \mathbf{x}$$

. . . ET EXEMPLES

Pour tout entier positif k, l'ensemble $\mathcal{E}_k$ est espace vectoriel. On a vu le cas de $\mathcal{E}_3$.

Les ensembles $\mathcal{F}$ et $\mathcal{G}$ de fonctions sont aussi des espaces vectoriels, mais pas $\mathcal{H}$.

L'ensemble $U = \{(x, 2x) \mid x \in \mathbb{R}\}$ est un espace vectoriel. Par contre, l'ensemble $W = \{(x, 2x+1) \mid x \in \mathbb{R}\}$ n'est pas un espace vectoriel: on a vu que les axiomes 1) et 2) ne sont pas valides.

Les matrices $\mathbb{M}_{2,3}$ forment une espace vectoriel; l'ensemble $\mathbb{M}_{m,n}$ est espace vectoriel pour tout entiers positifs m et n.

L'ensemble $\mathbb{P}_k$ est espace vectoriel, pour tout entier positif k.

NB: Un espace vectoriel a quatre ingrédients: l'ensemble de vecteurs, l'ensemble des scalaires, la définition d'addition, et la définition de multiplication scalaire. Ce n'est pas strictement correct de dire $\mathbb{R}^n$ est espace vectoriel sans mentionner le reste. On ne mentionne pas le reste seulement parce qu'il y a une addition et multiplication scalaire "standard" pour $\mathbb{R}^n$.

"sous"-espaces

On connait plusieurs espaces vectoriels: $\mathbb{R}^n$, équations, matrices, fonctions. On a vu aussi un exemple d'un espace vectoriel qui est *contenu* dans un autre. En particulier, plusieurs des axiomes étaient "hérités" de l'espace parent.

Notre but maintenant est de comprendre ces "sous"-espaces.

géometrie

On connait déja que $\mathbb{R}^n$ est un espace vectoriel. Rapellons maintenant l'exemple de

$$U = \{(x, 2x) \mid x \in \mathbb{R}\}.$$

On peut vérifier chacun des dix axiomes. C'est un exercice, mais, on commence a comprendre que c'est long!

Par contre, si on a déja vérifié chaque axiome pour $\mathbb{R}^2$ (ou même $\mathbb{R}^n$ directement!) on voit qu'on se répète. Par exemple, l'axiome 3) exige que:

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \qquad \text{pour tout } \mathbf{x}, \mathbf{y} \in U$$

Mais si $\mathbf{x}, \mathbf{y} \in U$, alors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Puisqu'on a déja vérifié l'axiome 3) pour *tout vecteurs dans* $\mathbb{R}^2$, c'est certainement valide pour tout vecteurs dans U. Un argument similaire s'applique à chaque axiome sauf $1), 2), 8), 9)$.

L'axiome 1) est valide puisque pour tout $\mathbf{x}, \mathbf{y} \in U$ on a

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 \end{bmatrix} + \begin{bmatrix} y_1 \\ 2y_1 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ 2x_1 + 2y_1 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ 2(x_1 + y_1) \end{bmatrix} \in U$$

L'axiome 2) est valide puisque pour tout $\mathbf{x} \in U$ et $t \in \mathbb{R}$ on a

$$t \cdot \mathbf{x} = t \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = t \cdot \begin{bmatrix} x_1 \\ 2x_1 \end{bmatrix} = \begin{bmatrix} tx_1 \\ t2x_1 \end{bmatrix} = \begin{bmatrix} tx_1 \\ 2(tx_1) \end{bmatrix} \in U$$

On pense que l'axiome 8) est valide car $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ joue le rôle de zéro pour tout $\mathbb{R}^2$. Il faut vérifier que $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ est dans U aussi, mais c'est facile car $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2(0) \end{bmatrix}$.

On pense que l'axiome 9) est valide car $-\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}$ joue le rôle d'inverse pour tout $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$. Il faut vérifier que $\begin{bmatrix} -x \\ -y \end{bmatrix}$ est dans U aussi, mais c'est facile car

$$\begin{bmatrix} x \\ y \end{bmatrix} \in U \iff y = 2x \iff -y = -2x \iff \begin{bmatrix} -x \\ -y \end{bmatrix} \in U$$

Note que il n'y a pas grand chose à vérifier pour 8) et 9)...

SOUS-ESPACES

Soit V un espace vectoriel: donc, les dix axiomes sont tous valides pour V. Soit U un sous-ensemble de V avec les mêmes opérations. Si U est aussi espace vectoriel, on dit que U est *sous-espace* de V. On cherche à savoir quand U est sous-espace.

Premièrement, on découvre quelques petits résultats en espaces vectoriels.

Question: Pourrait-on avoir deux vecteurs qui jouent le rôle de zéro? Pourrait-on avoir deux vecteurs qui jouent le rôle d'un inverse?

Mettons que $\mathbf{z}'$ et $\mathbf{z}''$ sont deux vecteurs qui jouent le rôle de zéro. Donc pour tout $\mathbf{x}$, on a $\mathbf{x} + \mathbf{z}' = \mathbf{z} = \mathbf{x} + \mathbf{z}''$. En particulier, puisque $\mathbf{z}'$ est un zéro on a

$$\mathbf{z}' + \mathbf{z}'' = \mathbf{z}''$$

et puisque $\mathbf{z}''$ est un zéro on a

$$\mathbf{z}' + \mathbf{z}'' = \mathbf{z}'$$

Donc $\mathbf{z}'' = \mathbf{z}' + \mathbf{z}'' = \mathbf{z}'$,

Soit $\mathbf{x}$ un vecteur. Mettons que $\mathbf{x}'$ et $\mathbf{x}''$ sont deux vecteurs qui jouent le rôle de $-\mathbf{x}$. Donc on a $\mathbf{x} + \mathbf{x}' = \mathbf{0} = \mathbf{x} + \mathbf{x}''$. Mais alors (grâce aux axiomes!) on a:

$$\mathbf{x}' = \mathbf{x}' + \mathbf{0} = \mathbf{x}' + (\mathbf{x} + \mathbf{x}'') = \mathbf{x}' + \mathbf{x} + \mathbf{x}'' = (\mathbf{x}' + \mathbf{x}) + \mathbf{x}'' = \mathbf{0} + \mathbf{x}'' = \mathbf{x}''$$

Donc on peut parler de "le" vecteur zéro, ainsi que "le" inverse de $\mathbf{x}$.

Proposition 1. *Soit V un espace vectoriel. Alors $\mathbf{0} = 0 \cdot \mathbf{x}$ pour tout $\mathbf{x}$ dans V.*

Si V est espace vectoriel, alors chaque axiome est valide pour V. Puisque V n'est pas vide, alors il existe un vecteur $\mathbf{x} \in V$. Donc:

$$0 \cdot \mathbf{x} = (0 + 0) \cdot \mathbf{x} = 0 \cdot \mathbf{x} + 0 \cdot \mathbf{x}$$

à cause de l'axiome 6). Mais grâce à l'axiome 8), $0 \cdot \mathbf{x}$ possède une inverse qu'on peut additionner aux deux côtés:

$$0 \cdot \mathbf{x} + (-0 \cdot \mathbf{x}) = 0 \cdot \mathbf{x} + 0 \cdot \mathbf{x} + (-0 \cdot \mathbf{x})$$

et donc:

$$\mathbf{0} = 0 \cdot \mathbf{x} + \mathbf{0}$$

qui veut dire que $0 \cdot \mathbf{x} = \mathbf{0}$.

Proposition 2. *Soit V un espace vectoriel. Alors $-\mathbf{x} = -1 \cdot \mathbf{x}$ pour tout $\mathbf{x}$ dans V.*

On applique les axiomes encore une fois

$$\mathbf{0} = 0 \cdot \mathbf{x} = (1 + (-1)) \cdot \mathbf{x} = 1 \cdot \mathbf{x} + (-1) \cdot \mathbf{x} = \mathbf{x} + (-1) \cdot \mathbf{x}$$

et on voit alors que $\mathbf{0} = \mathbf{x} + (-1) \cdot \mathbf{x}$, voulant dire que $(-1) \cdot \mathbf{x}$ est l'inverse de $\mathbf{x}$.

Retournons à la question: comment déterminer si U est sous-espace? C'est le test de sous-espace:

Théorème 3. *Soit V un espace vectoriel, et soit U un sous-ensemble de V, muni des mêmes opérations que V. Alors U est sous-espace de V si*

(1) U n'est pas vide
(2) $\mathbf{x} + \mathbf{y} \in U$ pour tout $\mathbf{x}, \mathbf{y} \in U$
(3) $t \cdot \mathbf{x} \in U$ pour tout $\mathbf{x} \in U$ et $t \in \mathbb{R}$

La vérification des axiomes $3), 4), 5), 6), 7), 10)$ est automatique. Les axiomes 1) et 2) font parti du test. Pourquoi est-ce qu'on n'a pas besoin de vérifier 8) et 9) aussi? Pour tout $\mathbf{x} \in U$, $0 \cdot \mathbf{x}$ et $-1 \cdot \mathbf{x}$ sont dans U aussi, grâce à l'axiome 2). Mais alors $0 \cdot \mathbf{x} = \mathbf{0} \in U$ et $-1 \cdot \mathbf{x} = -\mathbf{x} \in U$ pour tout $\mathbf{x} \in U$.

NB: On peut facilement montrer que U n'est pas vide en montrant que $\mathbf{0} \in U$. C'est utile, car si $\mathbf{0} \notin U$, on peut arrêter et conclure toute de suite que U n'est pas sous-espace. C'est la raison pour la deuxième version, équivalente:

Théorème 4. *Soit V un espace vectoriel, et soit U un sous-ensemble de V, muni des mêmes opérations que V. Alors U est sous-espace de V si*

(1) $\mathbf{0} \in U$
(2) $\mathbf{x} + \mathbf{y} \in U$ pour tout $\mathbf{x}, \mathbf{y} \in U$
(3) $t \cdot \mathbf{x} \in U$ pour tout $\mathbf{x} \in U$ et $t \in \mathbb{R}$

NB: Il faut vérifier les axiomes de fermeture pour U et non V. On connait déja que V est espace vectoriel, donc il s'agit de savoir si la somme de deux vecteurs dans U est un vecteur dans l'ensemble U.

EXEMPLES

Exemple: Est-ce que $U = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y = 2x \right\}$ est sous-espace de $\mathbb{R}^2$?

C'est l'exemple qu'on a vu ci-haut. Certainement U n'est pas vide. On vérifie que

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} \in U$$

Puisque $y_1 + y_2 = 2x_1 + 2x_2 = 2(x_1 + x_2)$, la condition est satisfaite et la somme est dans U. On vérifie aussi que

$$t \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} tx \\ ty \end{bmatrix} \in U$$

puisque $ty = t(2x) = 2(tx)$, et donc les deux axiomes de fermeture sont satisfaits.

Exemple: Est-ce que $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid y = 2x + 1 \right\}$ est sous-espace de $\mathbb{R}^2$?

Certainement W n'est pas vide. On calcule que

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}$$

mais ici $y_1 + y_2 = (2x_1 + 1) + (2x_2 + 1) = 2(x_1 + x_2) + 2 \neq 2(x_1 + x_2) + 1$. Donc l'axiome 1) n'est pas satisfaite. Comme alternative, il aurait suffit de donner un seul exemple: $\begin{bmatrix} 3 \\ 7 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix} \notin W$, car $8 \neq 2(3) + 1$. Comme troisième alternative, si on avait vérifié si $\mathbf{0} \in W$, on aurait trouvé que non.

Exemple: Soit $\mathbb{M}_{2,2}$ l'ensemble des matrices de taille 2×2. Est-ce que

$$D_2 = \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a, b \in \mathbb{R} \right\}$$

est sous-espace de $\mathbb{M}_{2,2}$?

On ne l'a pas dit, donc les opérations sont les "ordinaires". Note qu'il n'y a pas de contrainte: c'est plutôt la forme (diagonale) qui est la contrainte.

On voit que $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \in D_2$. On vérifie que

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} = \begin{bmatrix} a + c & 0 + 0 \\ 0 + 0 & b + d \end{bmatrix} = \begin{bmatrix} a + c & 0 \\ 0 & b + d \end{bmatrix} \in U$$

et que

$$t \cdot \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} ta & t0 \\ t0 & tb \end{bmatrix} = \begin{bmatrix} ta & 0 \\ 0 & tb \end{bmatrix} \in U$$

donc c'est un sous-espace de $\mathbb{M}_{2,2}$.

Exemple: Est-ce que $X = \left\{ \begin{bmatrix} a & b \\ c & 0 \end{bmatrix} \mid a, b, c \in \mathbb{R} \right\}$ est sous-espace de $\mathbb{M}_{2,2}$?

C'est largement similaire:

$$\begin{bmatrix} a & b \\ c & 0 \end{bmatrix} + \begin{bmatrix} d & e \\ f & 0 \end{bmatrix} = \begin{bmatrix} a + d & b + e \\ d + f & 0 + 0 \end{bmatrix} = \begin{bmatrix} a + d & b + e \\ d + f & 0 \end{bmatrix} \in X$$

et

$$t \cdot \begin{bmatrix} a & b \\ c & 0 \end{bmatrix} = \begin{bmatrix} ta & tb \\ tc & t0 \end{bmatrix} = \begin{bmatrix} ta & tb \\ tc & 0 \end{bmatrix} \in X$$

Exemple: Est-ce que $Y = \left\{ \begin{bmatrix} a & b \\ c & 1 \end{bmatrix} \mid a, b, c \in \mathbb{R} \right\}$ est sous-espace de $\mathbb{M}_{2,2}$?

C'est un peu différent:

$$\begin{bmatrix} a & b \\ c & 1 \end{bmatrix} + \begin{bmatrix} d & e \\ f & 1 \end{bmatrix} = \begin{bmatrix} a+d & b+e \\ d+f & 1+1 \end{bmatrix} = \begin{bmatrix} a+d & b+e \\ d+f & 2 \end{bmatrix} \notin Y$$

et

$$t \cdot \begin{bmatrix} a & b \\ c & 1 \end{bmatrix} = \begin{bmatrix} ta & tb \\ tc & t1 \end{bmatrix} = \begin{bmatrix} ta & tb \\ tc & t \end{bmatrix} \notin Y$$

En fait, si $t = 1$ la multiplication scalaire reste dans U... mais ça n'a pas d'importance: pour être valide, un axiome doit *toujours* être valide. On aurait pu aussi vérifier que $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \notin U$. On a démontré trois fois que Y n'est pas sous-espace!

Exemple: Soit $\mathbb{P}_2$ l'ensemble des polynômes de degré au plus 2. Est-ce que $\{ax^2 + bx + 1 \mid a, b \in \mathbb{R}\}$ est sous-espace de $\mathbb{P}_2$?

On ne l'a pas dit, donc les opérations sont les "ordinaires".

Certainement l'ensemble n'est pas vide. Le "1" nous donne des soupçons... En fait ce n'est pas un sous-espace. On voit que le zéro (c'est-à-dire $0x^2 + 0x + 0$) n'est pas dans l'ensemble, car la constante n'est pas 1. Alternativement, on voit que l'addition n'est pas fermé, car (par exemple) on a $(x^2 + 1) + (4x + 1) = x^2 + 4x + 2$ qui n'est pas dans l'ensemble. Alternativement on voit que la multiplication scalaire n'est pas fermé... un exercice pour vous!

Exercice: Qu'arrive si on change le "1" dans l'exemple précédent à "17"? à "0"?

Exemple: Est-ce que $Z = \{p(x) \in \mathbb{P}_2 \mid p(2) = 0\}$ est sous-espace?

Est-ce que Z est non-vide? Oui, car le polynôme $0x^2 + 0x + 0$ possède la bonne propriété: en substituant la valeur $x = 2$, on obtient $0(2)^2 + 0(2) + 0 = 0$.

Si $p(x)$ et $q(x)$ sont deux polynômes tel que $p(2) = q(2) = 0$, alors $p(2) + q(2) = 0 + 0 = 0$. On conclut que $p + q \in Z$. De plus $t \cdot p(2) = t(0) = 0$. On conclut que $t \cdot t \in Z$. Donc les axiomes de fermeture son valides.

Alternativement, $p(x) \in \mathbb{P}_2$ veut dire que $p(x) = ax^2+bx+c$. Donc $p(2) = 0$ veut dire que $a(2)^2+b(2)+c = 4a+2b+c = 0$. On voit que $Z = \{ax^2 + bx + c \mid 4a + 2b + c = 0\}$ est une forme équivalente. On vérifie

$$(ax^2 + bx + c) + (dx^2 + ex + f) = (a + d)x^2 + (b + e)x + (c + f) \in Z$$

car ... car pourquoi? C'est une petite exercice.

Exemple: Est-ce que $Z = \{p(x) \in \mathbb{P}_2 \mid p(2) = 7\}$ est sous-espace?

Certainement $x^2 + 2x - 1 \in Z$, car $(2)^2 + 2(2) - 1 = 7$. Mais $-3 \cdot (x^2 + 2x - 1) = -3x^2 - 6x + 3 \notin Z$, car $-3(2)^2 - 6(2) + 3 = -21 \neq 7$. Donc la multiplication scalaire n'est pas fermé, et ce n'est pas sous-espace.

UN COMMENTAIRE

Il y a essentiellement deux façons de donner un sous-espace: soit comme ensemble de vecteurs ayant une certaine forme (ex: matrices diagonales) ou ensemble de matrices qui satisfont une condition (ex: $y = 2x$). Retournez aux exemples ci-haut: lesquels sont lesquels? Mais est-ce vraiment deux façons différents? Je vous laisse avec deux exercices "complètements nouveaux"... que vous n'avez certainement jamais vu parmi les exemples précédents...

Exercice: Est-ce que $\left\{ \begin{bmatrix} x \\ 2x \end{bmatrix} \mid x \in \mathbb{R} \right\}$ est sous-espace de $\mathbb{R}^2$?

Exercice: Est-ce que $\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid b = 0, c = 0 \right\}$ est sous-espace de $\mathbb{M}_{2,2}$?

COMMENT DÉCRIRE UN ESPACE?

Commençons avec un exemple: la droite

$$L = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x + y = 0 \right\}.$$

Ce n'est nul autre que l'ensemble des points (x, y) (écrit en forme de vecteur en position standard) tel que $x = -y$. Sachant y, on connait x. Donc y fonctionne comme paramètre. Formellement, posons $y = t$ et on a

$$L = \left\{ \begin{bmatrix} -t \\ t \end{bmatrix} \mid t \in \mathbb{R} \right\} = \left\{ t \begin{bmatrix} -1 \\ 1 \end{bmatrix} \mid t \in \mathbb{R} \right\}.$$

On a la forme paramétrique de la droite, avec vecteur directeur $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ et passant par le point. . . ?

Exercice: Donner le point de départ de cette droite. Mieux: on connait que l'équation n'est pas unique, donc donner une *deuxième* possibilité pour le point de départ. Lequel te semble plus naturel?

On voit qu'on peut transformer une équation "normale" en forme "paramétrique". Peut-on faire l'inverse?

Exercice: Donner la forme "normale" pour les trois droites

$$L_1 = \left\{ c \begin{bmatrix} 3 \\ -0.5 \end{bmatrix} \mid c \in \mathbb{R} \right\}$$

$$L_2 = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} + q \begin{bmatrix} 4 \\ 2 \end{bmatrix} \mid q \in \mathbb{R} \right\}$$

$$L_3 = \left\{ \begin{bmatrix} \sqrt{2} \\ -2 \end{bmatrix} + a \begin{bmatrix} 3 \\ 7 \end{bmatrix} \mid a \in \mathbb{R} \right\}$$

Faire des graphqiues pour chacune, montrant bien le point de départ et le vecteur directeur.

Exercice: Lesquels de ces droites sont des espaces vectoriels? [indice: test de sous-espace]

On s'intéresse plus à celles qui sont des espaces vectoriels; une droite qui ne passe pas par l'origine ne l'est certainement pas [pourquoi?], donc on se limite à des droites qui ont la forme — pardon, *les* formes — suivantes:

$$\left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid ax + by = 0 \right\} = \left\{ t \begin{bmatrix} -b \\ a \end{bmatrix} \mid t \in \mathbb{R} \right\}$$

C'est exactement l'idée de "direction défendue" et de "direction permise".

Peut-on généraliser cette méthode à $\mathbb{R}^n$? Considérons

$$M = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + 3z = 0 \right\}.$$

En paramétrique on aura $x = -2y - 3z$, donc *deux* paramètres y et z. Si on pose $y = s$ et $z = t$ on a

$$M = \left\{ \begin{bmatrix} -2s - 3t \\ s \\ t \end{bmatrix} \mid s, t \in \mathbb{R} \right\} = \left\{ s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Une équation normale en $\mathbb{R}^2$ donne *deux* paramètres. On peut décrire cette sous-espace [est-ce vraiment un sous-espace?] en donnant les deux directions permises, c'est-à-dire, les deux vecteurs $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ et $\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$.

On dit que M est *engendré* par les vecteurs $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$. Afin de décrire M, il suffit de donner ces deux vecteurs.

Ce serait largement similaire pour une équation normale en $\mathbb{R}^n$.

Une dernière exemple, l'équation $x + z = 0$ en $\mathbb{R}^3$.

$$N = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + z = 0 \right\}.$$

On a $x = -z$, donc un seul paramètre? Non, puisque afin de déterminer le vecteur (x, y, z) il faut donner les trois valeurs x, y, et z. Il reste y. Mais y ne doit obéir à aucune condition, donc y fonctionne aussi comme paramètre. Proprement on met $y = s$ et $z = t$ afin d'obtenir

$$N = \left\{ \begin{bmatrix} -t \\ s \\ t \end{bmatrix} \mid s, t \in \mathbb{R} \right\} = \left\{ s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

ENGENDRANCE

On commence avec quelques définitions.

Définition 1. *Soit V un espace vectoriel, et $x_1, x_2, \dots, x_n \in V$.*

- *Si $a_1, a_2, \dots, a_n \in \mathbb{R}$, on dit que $x = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ est* combainaison linéaire *de $\{x_1, x_2, \dots, x_n\}$.*
- *L'ensemble de toutes les combinaisons linéaires de $\{x_1, x_2, \dots, x_n\}$ est dit* l'espace engendré par $\{x_1, x_2, \dots, x_n\}$, *et on écrit* $\operatorname{span}\{x_1, x_2, \dots, x_n\}$.
- *Un espace vectoriel W est* engendré par $\{x_1, x_2, \dots, x_n\}$ *si $W = \operatorname{span}\{x_1, x_2, \dots, x_n\}$.*

On voit que

$$\operatorname{span}\{x_1, x_2, \dots, x_n\} = \{a_1x_1 + a_2x_2 + \cdots + a_nx_n \mid a_1, \dots, a_n \in \mathbb{R}\}$$

Revenons aux examples.

$$M = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + 3z = 0 \right\} = \operatorname{span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Afin de simplifier la présentation, mettons $\mathbf{u} = \begin{bmatrix} -2 & 1 & 0 \end{bmatrix}^T$ et $\mathbf{v} = \begin{bmatrix} -3 & 1 & 0 \end{bmatrix}^T$.

Tout vecteur en M est combinaison linéaire de $\{\mathbf{u}, \mathbf{v}\}$, et toute combinaison linéaire de $\{\mathbf{u}, \mathbf{v}\}$ est vecteur en M. On peut vérifier ceci directement. Tout vecteur en M a la forme $\begin{bmatrix} -2y - 3z & y & z \end{bmatrix}^T$ et on a donc

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2y - 3z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

Ce n'est pas que des vecteurs géométriques. On peut parler d'engendrance dans n'importe quel espace vectoriel.

$$U = \left\{ ax^2 + bx + c \mid c = 2a \right\} = \operatorname{span}\left\{ x^2 + 2, x \right\}$$

Pourquoi est-ce valide? Un vecteur arbitraire dans U est de la forme $ax^2 + bx + 2a$ (c'est ce que veut dire la condition $c = 2a$). Donc

$$ax^2 + bx + c = ax^2 + bx + 2a = ax^2 + 2a + bx = a(x^2 + 2) + b(x)$$

On vient de montrer que tout vecteur en U peut s'écrire comme combinaison linéaire de $x^2 + 2$ et x, ou plutôt que $U = \operatorname{span}\{x^2 + 2, x\}$.

NB: Ce n'est pas vrai, on n'a que montré que $U \subseteq \operatorname{span}\{x^2 + 2, x\}$. Peut-on aussi montrer que $\operatorname{span}\{x^2 + 2, x\} \subseteq U$?

On a dit que span $\{\cdots\}$ est *l'espace* engendré. Est-ce vraiment une espace? Oui.

Théorème 2. *Soit V un espace vectoriel.*

- *Si $x_1, x_2, \ldots, x_n \in V$, alors* span $\{x_1, x_2, \ldots, x_n\}$ *est un sous-espace de V.*
- *Si W un sous-espace quelconque de V, avec $\{x_1, x_2, \ldots, x_n\} \subseteq W$, alors* span $\{x_1, x_2, \ldots, x_n\} \subseteq W$.

Note que span $\{x_1, x_2, \ldots, x_n\}$ est *sous*-espace de V. C'est certainement un espace vectoriel, mais pas nécessairement V!

La démonstration de ce théorème repose entièrement sur le test de sous-espace. C'est-à-dire, il faut vérifier que span $\{x_1, x_2, \ldots, x_n\}$ satisfait les deux axiomes de fermeture. On vérifie un; l'autre est une exercice! Prennons deux éléments arbitraires dans l'espace engendré; en autres mots, deux combinaisons linéaires de $\{x_1, x_2, \ldots, x_n\}$.

$$\begin{aligned}(a_1x_1 + a_2x_2 + \cdots + a_nx_n) \ &+ \ (b_1x_1 + b_2x_2 + \cdots + b_nx_n)\\ &= (a_1 + b_1)x_1 + (a_2 + b_2)x_2 + \cdots + (a_n + b_n)x_n\end{aligned}$$

Puisque chaque a_i+b_i est un chiffre réel, la somme est aussi une combinaison linéaire de $\{x_1, x_2, \ldots, x_n\}$, et donc la somme est élément de span $\{x_1, x_2, \ldots, x_n\}$.

Question: Dans l'équation ci-haut, lesquels des symboles "+" sont des additions de *vecteurs*?

Voyons un exemple. Est-ce que l'ensemble des matrices diagonales de taille 2×2 est un espace vectoriel? On a déja vu cette question, mais on a maintenant une autre approche.

$$\begin{aligned}D_2 &= \left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \mid a, b \in \mathbb{R} \right\}\\ &= \left\{ \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix} \mid a, b \in \mathbb{R} \right\}\\ &= \left\{ a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \mid a, b \in \mathbb{R} \right\}\\ &= \operatorname{span} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}\end{aligned}$$

On a montré que l'ensemble D_2 est l'espace engendré par quelque chose (et on a déterminé le "quelque chose"). Donc grâce à notre théorème, D_2 est un espace vectoriel; même plus, un sous-espace de $\mathbb{M}_{2\times 2}$.

On comprend maintenant deux facons de d'écrire un espace vectoriel: une forme générale avec des conditions, ou comme un espace engendré par un ensemble de vecteurs. Tout espace peut s'écrire des deux manières: c'est exactement l'exemple du début de ce chapitre.

Il y a des avantages et désavantages pour chaque. Supposons qu'on cherche à savoir si un vecteur $\mathbf{x}$ est dans un espace vectoriel U. Prennons un exemple précis. Est-ce que le vecteur

$$\mathbf{x} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$$

est dans l'espace vectoriel donné dans deux formes différents:

$$M = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + 3z = 0 \right\} = \operatorname{span} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Essayons de répondre à la question pour chaque forme. Sachant la première forme, on a une méthode directe. Le vecteur $\mathbf{x}$ est élément de $\mathbb{R}^3$, mais la condition n'est pas satisfaite, car

$$x + 2y + 3z = (5) + 2(-3) + 3(1) = 2 \neq 0$$

Donc $\mathbf{x} \notin M$. Mais qu'arrive-t-il si on n'a que la deuxième forme? Est ce que $\mathbf{x}$ est dans l'espace engendré? Si oui, $\mathbf{x}$ serait combinaison linéaire des deux vecteurs, et on aura une solution à la suivante:

$$\begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

Ceci donne le système

$$\begin{aligned} 5 &= -2s - 3t \\ -3 &= s \\ 1 &= t \end{aligned}$$

On vous laisse l'exercice de montrer que ce système n'a aucune solution. Donc on redécouvre que $\mathbf{x} \notin M$.

NB: On voit encore une fois un système d'équations linéaires. On voudrait bien un beau jour avoir une méthode efficace de résoudre des tels systèmes sans se fier à ce que "les chiffres seront toujours faciles"...

Il y a une difficulté qui se présente avec des espaces présentés comme espaces engendrés. Quelle est la relation entre les suivants:

$$U = \operatorname{span}\{(1, 2, 0), (3, 1, 1)\} \qquad W = \operatorname{span}\{(0, 5, -1), (5, 0, 2)\}$$

Réponse: les deux espaces sont identiques. Inquiètant, vu qu'ils paraissent complètement différents. Voyons.

$$\begin{aligned}(0,5,-1) &= 3(1,2,0)-(3,1,1)\\(5,0,2) &= -(1,2,0)+2(3,1,1)\end{aligned}$$

Donc $(0,5,-1)$ et $(5,0,2)$ sont des éléments dans U. Donc l'espace qu'ils engendrent, c'est-à-dire W, fait partie de U: c'est notre théorème. Mais on a aussi

$$\begin{aligned}(1,2,0) &= .4(0,5,-1)+0.2(5,0,2)\\(3,1,1) &= .2(0,5,-1)+.6(5,0,2)\end{aligned}$$

Donc $(1,2,0)$ et $(3,1,1)$ sont des éléments dans W. Donc l'espace qu'ils engendrent, c'est-à-dire U, fait partie de W. On a $W \subseteq U$ et $U \subseteq W$, donc $U = W$.

On a pire.

Exercice: Montrer que U et V sont des espaces identiques:

$$U = \operatorname{span}\{(1,2,0),(3,1,1)\} \qquad V = \operatorname{span}\{(0,5,-1),(5,0,2),(10,10,2)\}$$

SOUS-ESPACES DE $\mathbb{R}^n$

Comme application de l'engendrance, on détermine tous les sous-espaces de $\mathbb{R}^n$.

Commençons avec $\mathbb{R}^2$. L'ensemble $\{\mathbf{0}\}$ et l'ensemble $\mathbb{R}^2$ sont des sous-espaces (test de sous-espace!). Est-ce qu'il y a d'autres? Posons U un sous-espace avec $\{\mathbf{0}\} \neq U \neq \mathbb{R}^2$. Donc, il existe un vecteur $\mathbf{x} \in U$ avec $\mathbf{x} \neq \mathbf{0}$. Donc $\operatorname{span}\{\mathbf{x}\} \subseteq U$. Mais $\operatorname{span}\{\mathbf{x}\}$ est exactement une droite passant par l'origine (indice: revoir l'exemple au début de ce chapitre). Donc les droites passant par l'origine sont des sous-espaces. Est-ce qu'il y en a d'autres? Posons V un sous-espace qui contient un vecteur $\mathbf{x} \neq \mathbf{0}$ mais qui n'est pas une droite. Donc V contient un autre vecteur $\mathbf{y}$ qui n'est pas parallèle à $\mathbf{x}$. Mais alors $\operatorname{span}\{\mathbf{x},\mathbf{y}\}$ serait $\mathbb{R}^2$ et grâce à notre théorème, $\mathbb{R}^2 = \operatorname{span}\{\mathbf{x},\mathbf{y}\} \subseteq V \subseteq \mathbb{R}^2$, donc $V = \mathbb{R}^2$.

Un dessin ce montre très utile ici...

Résultat final: les seuls sous-espaces de $\mathbb{R}^2$ sont $\{\mathbf{0}\}$, $\mathbb{R}^2$, et les droites passant par l'origine.

Un argument similaire montre que pour $\mathbb{R}^3$, les seuls sous-espaces sont $\{\mathbf{0}\}$, $\mathbb{R}^3$, les droites passant par l'origine et les plans passant par l'origine.

En $\mathbb{R}^4$, on aura les sous-espaces $\{\mathbf{0}\}$, $\mathbb{R}^3$, les droites passant par l'origine, les plans passant par l'origine et les "hyperplans" passant par l'origine.

Une droite passant par l'origine est un espace engendré par un seul vecteur non-nul. Un plan passant par l'origine est un espace engendré par deux vecteurs qui ne sont pas parallèles. Un "hyperplan" est un espace engendré par trois vecteurs qui ne sont pas *coplanaires*.

Comment généraliser? On a besoin d'un nouveau concept: *l'indépendence linéaire*.

UN PROBLÈME

On se souvient que les deux espaces suivants sont en fait identiques:

$$U = \operatorname{span}\left\{ \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}2\\1\\1\end{bmatrix} \right\} \qquad W = \operatorname{span}\left\{ \begin{bmatrix}0\\5\\-1\end{bmatrix}, \begin{bmatrix}5\\0\\2\end{bmatrix}, \begin{bmatrix}10\\10\\2\end{bmatrix} \right\}$$

On observe aussi (avec un peu de papier brouillon) que $(10, 10, 2)$ est combinaison linéaire de $(0, 5, -1)$ et $(5, 0, 2)$. Ajouter un vecteur qui est combinaison linéaire de ceux qu'on a deja ne semble pas changer l'espace engendré. Un autre exemple:

$$U = \operatorname{span}\left\{ \begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}2\\1\\1\end{bmatrix} \right\} \qquad Z = \operatorname{span}\left\{ \begin{bmatrix}0\\5\\-1\end{bmatrix}, \begin{bmatrix}5\\0\\2\end{bmatrix}, \begin{bmatrix}10\\10\\2\end{bmatrix}, \begin{bmatrix}5\\5\\1\end{bmatrix}, \begin{bmatrix}10\\0\\3\end{bmatrix} \right\}$$

Exercice: Montrer que les deux nouveaux vecteurs sont combinaison linéaire de ceux dans W; aussi montrer que $U = Z$.

Il y a du gaspillage: pourquoi donner une liste de trois ou cinq vecteurs comme ensemble engendrant, lorsqu'une liste de deux vecteurs suffit? Ici, c'est un exemple petit, et il y a plusieurs "trucs" pour voir que l'ensemble de Z n'est pas optimal. Mais en générale, que faire? Comment trouver un ensemble engendrant *minimal*?

On soupçonne que le problème est précisément que dans l'ensemble engendrant, il y a des vecteurs qui sont combinaison linéaire des autres.

Exercice: Si $\mathbf{z} \in \operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$ et chaque $\mathbf{y}_i \in \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$, montrer que $\mathbf{z} \in \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$. Expliquer pourquoi l'exercice précédente est un cas spécial de ceci.

DÉPENDENCE ET INDÉPENDENCE

Définition 1. *Un ensemble de vecteurs* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_n\}$ *est* linéairement dépendent *si il existe des scalaires* $a_1, a_2, \ldots, a_{n-1}, a_n$ *qui ne sont pas tous zéro tel que* $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} + a_n\mathbf{x}_n = \mathbf{0}$

L'idée clé est "pas tous zéro". Il se peut que certains des a_i soient zéro mais *pas tous*. Pourquoi est-ce important? Si les a_i ne sont pas tous zéro, il y a au moins un qui n'est pas zéro. Par exemple, disons $a_n \neq 0$. Alors:

$$\begin{aligned} & a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} + a_n\mathbf{x}_n = \mathbf{0} \\ \iff & a_n\mathbf{x}_n = -a_1\mathbf{x}_1 - a_2\mathbf{x}_2 - \cdots - a_{n-1}\mathbf{x}_{n-1} \\ \iff & \mathbf{x}_n = -\frac{a_1}{a_n}\mathbf{x}_1 - \frac{a_2}{a_n}\mathbf{x}_2 - \cdots - \frac{a_{n-1}}{a_n}\mathbf{x}_{n-1} \end{aligned}$$

La dernière égalité est valide parce que $a_n \neq 0$. Il semble que si on a une dépendence on peut écrire un des vecteurs comme combinaison linéaire des autres. Ce n'est peut-être pas valide pour $\mathbf{x}_n$, mais c'est valide pour au moins un $\mathbf{x}_i$.

On dit parfois qu'une équation $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} + a_n\mathbf{x}_n = \mathbf{0}$ avec pas tous les a_i zéro est une *relation de dépendance*.

Définition 2. *Un ensemble de vecteurs* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_n\}$ *est* linéairement indépendent *si la seule solution à* $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} + a_n\mathbf{x}_n = \mathbf{0}$ *est* $a_1 = a_2 = \cdots = a_{n-1} = a_n = 0$.

L'idée clé est "la seule solution". On n'a aucun autre choix que de mettre chaque $a_i = 0$. On sait alors que *chaque* $\mathbf{x}_i$ n'est pas dans l'espace engendré par les autres. Par exemple, si $\mathbf{x}_n$ est dans span $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}\}$ alors on aura des chiffres réels $a_1, a_2, \ldots, a_{n-1}$ tel que

$$\begin{aligned} & \mathbf{x}_n = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} \\ \iff & a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} - 1\mathbf{x}_n = \mathbf{0} \end{aligned}$$

les coefficients ne sont pas tous zéro, car $a_n = -1 \neq 0$ donc l'ensemble serait linéairement dépendant. Mais on a dit que c'est linéairement indépendant, donc $\mathbf{x}_n$ n'est pas dans span $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}\}$.

Ces deux définitions sont équivalents: "être linéairement indépendant" c'est exactement la même chose que "ne pas être linéairement dépendant".

Exemple: Est-ce que $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \end{bmatrix} \right\}$ est dépendant ou indépendant?

On tente de résoudre l'équation

$$r \begin{bmatrix} 1 \\ 2 \end{bmatrix} + s \begin{bmatrix} 1 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Ceci donne:

$$r + s + t = 0 \qquad 2r + 3s + 4t = 0$$

En éliminant le r on obtient $s + 2t = 0$; substituant dans la première on obtient $r - t = 0$. Donc on a $s = -2t$ et $r = t$, où t parait être paramètre.[1] Puisque t est

[1] Encore une fois on cherche une méthode systématique et efficace de résoudre des systèmes linéaires!

paramètre, on peut la choisir n'importe comment. En particulier, on peut la choisir *pas égal à zéro*, par exemple $t = 1$. Donc $s = -2$ et $r = 1$. Peut importe r, s, on sait que les valeurs ne sont *pas tous zéro* car $t \neq 0$.

$$1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Or, on a trouver une solution ou les coefficients ne sont pas tous zéro. C'est une relation de dépendence. Les trois vecteurs sont linéairement dépendant.

NB: Chaque choix de $t \neq 0$ aurait donné une relation dépendance.

Exemple: Est-ce que $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ est dépendant ou indépendant?

On tente de résoudre

$$a \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Ceci donne:

$$a + b + c = 0 \qquad a + b = 0 \qquad a = 0$$

La troisième équation donne $a = 0$, et alors la deuxième donne $b = 0$, et alors la première donne $a = 0$. Donc la seule solution est $a = b = 0$. L'ensemble est indépendent.

Comment savoir si un ensemble est dépendant ou indépendant? On forme une combinaison linéaire arbitraire, et de solutionner pour les coefficients. Si les coefficients sont *obligatoirement tous zéro* alors l'ensemble est dépendant; s'il existe une solution ou les coefficients ne sont pas tous zéro alors l'ensemble est indépendant.

UN PEU DE THÉORIE

Théorème 3. *Soit* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *un ensemble dépendent. Alors* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{y}\}$ *est dépendent pour tout vecteur* $\mathbf{y}$.

Si on a $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = \mathbf{0}$ avec au moins un $a_i \neq 0$, alors on a $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n + 0\mathbf{y} = \mathbf{0}$ avec au moins un coefficient non-zéro (le a_i!).

Théorème 4. *Soit* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_n\}$ *un ensemble indépendent. Alors* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}\}$ *est indépendent.*

C'est équivalent au théorème précédente. Si on avait une relation de dépendence $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} = \mathbf{0}$ on aurait une relation de dépendance $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_{n-1}\mathbf{x}_{n-1} + 0\mathbf{x}_n = \mathbf{0}$.

NB: On conclut que si on veut éliminer la dépendence, il faudrait *enlever* des vecteurs et non en ajouter. Si on veut éliminer l'indépendance il faudrait *ajouter* des vecteurs et non en enlever.

On peut découvrir certains petits résultats.

Proposition 5. $\{\mathbf{0}\}$ *est dépendent*

On voit ceci directement, car $1 \cdot \mathbf{0} = \mathbf{0}$. À gauche, on a une combinaison linéaire où les coefficients ne sont pas tous zéro, donc c'est un relation de dépendence.

Proposition 6. *Tout ensemble contenant le vecteur* $\mathbf{0}$ *est dépendant.*

L'ensemble $\{\mathbf{0}\}$ est dépendant, et ceci ne change pas si on ajoute d'autre vecteurs.

Proposition 7. $\{\mathbf{x}, \mathbf{y}\}$ *est dépendant si et seulement si* $\mathbf{x} = t\mathbf{y}$ *ou* $\mathbf{y} = t\mathbf{x}$ *pour un* $t \in \mathbb{R}$.

L'ensemble $\{\mathbf{x}, \mathbf{y}\}$ est dépendent exactement lorsqu'il existe des scalaires a, b tel que $a\mathbf{x} + b\mathbf{y} = \mathbf{0}$, avec soit $a \neq 0$ ou $b \neq 0$. Donc on aura soit $\mathbf{x} = \frac{b}{a}\mathbf{y}$ ou $\mathbf{y} = \frac{a}{b}\mathbf{x}$.

Question: Si $a = 0$ et $b \neq 0$, qu'est-ce qui se passe? Et si $\mathbf{x} = \mathbf{0}$?

Quelques exemples

Exemple: Est-ce que $\{x^3 + 17x^2 + x + 1, 5x^2 - 12x + 15\}$ dans $\mathbb{P}_3$ est indépendant?

Il s'agit de deux vecteurs. On voit que l'un n'est pas multiple de l'autre et donc ils sont indépendants.

Exemple: Est-ce que $\{x^3 + 17x^2 + x + 1, 0x^3 + 0x^2 + 0x + 0\}$ dans $\mathbb{P}_3$ est indépendant? Il s'agit de deux vecteurs. On voit que l'un est multiple de l'autre: $(0x^3 + 0x^2 + 0x + 0) = 0(x^3 + 17x^2 + x + 1)$ et donc ils sont dépendants. Alternativement, on voit que $\mathbf{0}$ est dans l'ensemble, alors c'est dépendant.

Exemple: Est-ce que $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\}$ est indépendant?

On pose

$$r\begin{bmatrix}1\\2\\3\end{bmatrix}+s\begin{bmatrix}4\\5\\6\end{bmatrix}+t\begin{bmatrix}7\\8\\9\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

pour obtenir

$$r+4s+7t=0 \qquad 2r+5s+8t=0 \qquad 3r+6s+9t=0$$

Éliminant r entre les deux premières on obtient $3s+6t=0$ ou $s=-2t$. Substituant dans les trois équations on obtient:

$$r-t=0 \qquad 2r-2t=0 \qquad 3r-3t=0$$

Donc t est paramètre et $s=-2t, r=t$. On choisit $t\neq 0$, par exemple $t=1$ pour obtenir $r=1, s=-2, t=1$ et la relation de dépendance

$$1\begin{bmatrix}1\\2\\3\end{bmatrix}-2\begin{bmatrix}4\\5\\6\end{bmatrix}+1\begin{bmatrix}7\\8\\9\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

Suivant les commentaires après la définition, on peut récrire comme:

$$\begin{bmatrix}1\\2\\3\end{bmatrix}=2\begin{bmatrix}4\\5\\6\end{bmatrix}-1\begin{bmatrix}7\\8\\9\end{bmatrix}$$

montrant que les premier vecteur est dans l'espace engendré par les deux autres.

Exemple: Est-ce que $\{x^3+1, x^2+1, x^3+x^2\}$ est linéairement dépendent?

On demande la dépendance et non l'indépendance, mais la méthode est pareil. On pose

$$r(x^3+1)+s(x^2+1)+t(x^3+x^2)=\mathbf{0}=0x^3+0x^2+0x+0$$

Ceci donne

$$(r+t)x^3+(s+t)x^2+(0)x+(r+s)=\mathbf{0}=0x^3+0x^2+0x+0$$

et donc

$$r+t=0 \qquad s+t=0 \qquad r+s=0$$

Les deux premières équations donnent $r=s$; les deux dernières donnent $r=t$. Donc $r=s=t$ et donc $r+t=2r=0$ implique que $r=s=t=0$.

Il n'y a que la solution "tous zéro", donc les polynômes sont indépendants.

Exemple: Est-ce que $\{\sin(x), \cos(x), x\}$ est linéairement indépendent?

On pose

$$a(\sin(x))+b(\cos(x))+c(x)=\mathbf{0}$$

Ici, il n'y a pas un système évident. Par contre, il faut que cette équation soit valide *pour tout* x. On met $x=0,\pi,\pi/2$ pour donner

$$a(0)+b(1)+c(0)=0a(0)+b(-1)+c(-\pi) \qquad =0a(1)+b(0)+c(\pi/2)=0$$

On conclut que $b = 0$, que $c = b/(-\pi) = 0$, et $a = -c\pi/2 = 0$. Il n'y a que la solution "tous zéro", donc les fonctions sont indépendants.

NB: Ce n'est pas évident quelles valeurs choisir pour x.

Exemple: Est-ce que $\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \right\}$ est linéairement indépendent?

Encore une fois on met

$$a \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

pour obtenir

$$a + b + c = 0 \qquad a + b + c = 0 \qquad d = 0 \qquad a + 2c + d = 0$$

C'est bizarre: on a la même équation deux fois! On voit quand même que $d = 0$. Les deux autres équations donnent $b - c = 0$. Alors $b = c$ et $a = -2c$. On a plusieurs relations de dépendance, en choisissnat n'importe quelle valeur non-zéro pour c. Par exemple $a = 2, b = -1, c = -1, d = 0$.

On obtient alors la relation de dépendance

$$2 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 1 \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} - 1 \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} + 0 \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Suivant les commentaires après la définition on peut écrire:

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = -2 \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} + 1 \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$$

montrant que

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \right\}.$$

On aurait pu également montré que

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in \text{span} \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \right\}. \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \qquad \in \text{span} \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \right\}.$$

Exemple: Montrer que

$$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \right\}.$$

Il y a au moins deux façons de faire ceci. On peut tenter d'écrire le vecteur comme combinaison linéaire et observer qu'il n'y a aucune solution. Ou on peut regarder en bas à gauche...

On voit que $\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \right\}$ est un ensemble linéairement dépendant, et il est donc possible d'écrire au moins un des vecteurs comme combinaison linéarie des autres. Par contre, il n'est pas possible de faire ceci pour *tout* vecteur. C'est le "au moins un" dans la définition de la dépendance.

MAT 1741		8. Engendrance et Indépendance
Mike Newman	Université d'Ottawa	automne 2008

ENGENDRANCE ET INDÉPENDANCE

On se rapelle les défintions.

Définition 1. *Un ensemble de vecteurs* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *engendre* V *si tout vecteur en* V *s'écrit comme combinaison linéaire; c'est-à-dire que pour chaque* $\mathbf{v} \in V$ *il existe des scalaires* $a_1, a_2, \ldots, a_n$ *tel que* $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = \mathbf{v}$.

Alternativement, pour tout $\mathbf{v} \in V$, l'équation $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = \mathbf{v}$ possède au moins une solution. On voit que l'engendrance est relié à l'existence.

Définition 2. *Un ensemble de vecteurs* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est linéairement dépendant dans* V *si la seule solution à* $a_1+a_2+\cdots+a_{\mathbf{x}-1}+a_{\mathbf{x}}n = \mathbf{0}$ *est* $a_1 = a_2 = \cdots = a_n = 0$.

Alternativement, l'équation $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = \mathbf{0}$ possède une solution unique. On voit que l'indépendance est relié à l'unicité.

Réduire

Mettons qu'on a un ensemble que engendre un espace vectoriel V. Est-ce un ensemble *minimal*? Ici, minimal veut dire que si on enlève un vecteur, l'ensemble n'engendrait plus V: tout vecteur est essentiel. Le résultat suivant aide:

Proposition 3. *L'ensemble* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-1}, \mathbf{x}_n\}$ *est dépendent dans* V *si et seulement si un des* x_i *est engendré par les autres.*

On a déja vu ce résultat (c'est essentiellement la définition de la dépendance) mais maintenant on veut l'appliquer. Un ensemble engendrant pour V *non-minimal* est alors exactement un ensemble engendrant pour V qui est *dépendant*. De plus, la dépendance indique exactement comment obtenir un ensemble engendrant plus petit: on enlève un des vecteurs dépendants.

Exemple: Est-ce que $\left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\2\end{bmatrix} \right\}$ engendre $\mathbb{R}^3$?

Pour répondre, on pose

$$\begin{bmatrix}x\\y\end{bmatrix} = a\begin{bmatrix}1\\2\end{bmatrix} + b\begin{bmatrix}1\\0\end{bmatrix} + c\begin{bmatrix}1\\1\end{bmatrix} + d\begin{bmatrix}2\\2\end{bmatrix}$$

ce qui donne

$$x = a + b + c + 2d \qquad y = 2a + c + 2d$$

On cherche une solution pour a, b, c, d. Encore un système linéaire! On verra bientôt l'élimination de Gauss-Jordan, mais pour le moment... Si on met (par exemple) $c = y$ et $b = x - y$, et $a = d = 0$ on voit qu'on a une solution:

$$x = (0) + (x - y) + (y) + 2(0) \qquad y = 2(0) + (y) + 2(0)$$

Donc pour un vecteur arbitraire dans $\mathbb{R}^2$, on peut l'écrire comme combinaison linéaire des vecteurs donnés. Vous pouvez probablement trouver *d'autres* solutions: c'est une indication qu'il y a une dépenance...

Exemple: Trouver un ensemble engendrant pour $\mathbb{R}^2$ de taille minimale dans

$$\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$$

On commencerait en vérifiant que l'ensemble engendre $\mathbb{R}^2$, car autrement la question serait impossible! C'est l'exemple précédent, donc on peut procéder. Il s'agit de déterminer si on peut enlever des vecteurs.[1]

On cherche premièrement une dépendance, c'est-à-dire une solution non-nulle[2] à:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \end{bmatrix} + d \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

ou en équations:

$$0 = a + b + c + 2d \qquad 0 = 2a + c + 2d$$

On a l'embarras du choix! Prenons $a = 1$, $b = 1$, $c = -2$. En autre mots, on a trouvé la relation de dépendance:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Note que les coefficients ne sont pas tous zéro, même si $d = 0$. Ceci permet d'écrire:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = -1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Donc $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ est engendré par les autres: c'est exactement la proposition ci-haut. Alors on peut l'enlever:

$$\mathbb{R}^2 = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\} = \text{span} \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$$

On a réduit l'ensemble, sans changer l'espace engendré! Le vecteur qu'on a enlevé ne donnait rien de *nouveau*.

[1] Si ce n'était pas engendrant pour $\mathbb{R}^2$, on ne pourrait certainement pas trouver un ensemble engendrant en enlevant des vecteurs!

[2] voulant dire une solution autre que "tous zéro"

Maintenant on a un "meilleur" ensemble, mais est-ce minimal? Sinon, on aurais une dépendance. Voyons:

$$\begin{bmatrix}0\\0\end{bmatrix} = r\begin{bmatrix}1\\2\end{bmatrix} + s\begin{bmatrix}1\\1\end{bmatrix} + t\begin{bmatrix}2\\2\end{bmatrix}$$

ou en équations

$$0 = r + s + 2t \qquad 0 = 2r + s + 2t$$

Pour résoudre, on soustrait la première de la deuxième pour donner $r = 0$. Donc on a maintenant $(0) + s + 2t = 0$ et $2(0) + s + 2t = 0$, en autres mots $s = -2t$. On doit prendre $r = 0$, mais on peut choisir, par exemple $t = 1$ et $s = -2$. On a une solution non-nulle (même si une des variables est forcée d'être zéro, elles ne sont pas toutes). La relation est:

$$\begin{bmatrix}0\\0\end{bmatrix} = 0\begin{bmatrix}1\\2\end{bmatrix} - 2\begin{bmatrix}1\\1\end{bmatrix} + 1\begin{bmatrix}2\\2\end{bmatrix}$$

Ou plutôt

$$\begin{bmatrix}2\\2\end{bmatrix} = 2\begin{bmatrix}1\\1\end{bmatrix}$$

Donc $\begin{bmatrix}2\\2\end{bmatrix}$ est engendré par les autres. Alors on peut l'enlever:

$$\mathbb{R}^2 = \operatorname{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\} = \operatorname{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$$

On a encore réduit l'ensemble sans changer l'espace engendré!

Est-ce minimal maintenant? On vérifie si on a une dépendance:

$$\begin{bmatrix}0\\0\end{bmatrix} = p\begin{bmatrix}1\\2\end{bmatrix} + q\begin{bmatrix}1\\1\end{bmatrix}$$

ou en équations

$$0 = p + q \qquad 0 = 2p + q$$

En soustrayant la première de la deuxième, on obtient $p = 0$, qui donne alors $q = 0$. Il n'y a aucune solution non-nulle, donc l'ensemble est indépendent. La proposition ci-haut affirme qu'aucun des vecteurs n'est engendré par les autres. Si on enlève un vecteur, on perd quelque chose: l'espace engendré diminuerait. Donc un ensemble engendrant minimal pour $\mathbb{R}^2$ est

$$\left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$$

NB: J'ai fait des choix essentiellement arbitraires. Par exemple, il existe plusieurs relations de dépendance à chaque étape, et de plus, il y eu un choix de quel vecteur à enlever. En faisant d'autres choix, vous pouvez découvrir d'autres ensembles minimales. Par contre on aurait chaque fois deux vecteurs.

C'était un exemple long, mais on a découvert un théorème:

Théorème 4. *Un ensemble engendrant minimal pour V est indépendant.*

Ce n'est rien d'autre que la proposition ci-haut dit de façon différent: c'est minimal parce qu'on ne peut plus rien enlever, parce qu'il n'y a pas de dépendance, parce que l'ensemble est indépendant.

Pour résumer, on a l'algorithme suivant.

Algorithme 5. *Soit $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ un ensemble de vecteurs dans V qui engendre V.*

On peut trouver un ensemble engendrant minimal contenu dans cet ensemble en suivant les étapes suivantes.

(0) Vérifier si l'ensemble engendre V. Si non, ARRÊTER, la tâche est impossible.

On répète alors les suivants:

(1) Vérifier si l'ensemble est indépendant. Si oui, ARRÊTER, l'ensemble est engendrant minimal.
(2) Si l'ensemble est dépendant, on peut écrire au moins un des vecteurs comme combinaison linéaire des autres. Enlever ce vecteur de l'ensemble.
(3) Retourner à l'étape (1).

Augmenter

L'engendrance et l'indépendance sont complémentaires. Ce qui nous mène à poser une question complémentaire.

Mettons qu'on a un ensemble indépendant dans un espace vectoriel V. Est-ce maximal? Ici, maximal veut dire que si on ajoute un vecteur l'ensemble ne serait plus indépendant. On a un résultat similaire à la proposition ci-haut:

Proposition 6. *L'ensemble $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ n'engendre pas V si et seulement si il existe un $\mathbf{y} \in V$ tel que $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n, \mathbf{y}\}$ est indépendant.*

On a déja vu ce résultat aussi (c'est essentiellement la définition de l'indépendance), c'est l'application maintenant. Un ensemble indépendant dans V *non-maximal* est exactement un ensemble indépendant qui *n'engendre pas V*. De plus la preuve de non-engendrance est un vecteur $\mathbf{y}$ pas engendré par les $\mathbf{x}_i$: c'est exactement le vecteur qu'on voudrait ajouter à l'ensemble.

Exemple: Est-ce que $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}$ est indépendant dans $\mathbb{M}_{2,2}$?

Pour répondre, on pose

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = s \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + t \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

et on obtient

$$0 = s + t \qquad 0 = 0 \qquad 0 = t \qquad 0 = s$$

On voit que la seule solution est $s = t = 0$, donc aucune solution non-nulle: l'ensemble est indépendant.

Exemple: Trouver un ensemble indépendant dant $\mathbb{M}_{2,2}$ maximal qui contient

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

On commencerait en vérifiant que l'ensemble est indépendant, car autrement la question serait impossible! C'est l'exemple précédent, donc on peut procéder. Il s'agit de déterminer si on peut ajouter des vecteurs.[3]

On cherche premièrement un vecteur qui n'est pas engendré par l'ensemble. C'est-à-dire, on cherche un vecteur qui ne peut *pas* s'exprimer comme combinaison linéaire suivante:

$$a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ a & b \end{bmatrix}$$

Un exemple[4] d'une matrice qui n'est pas de cette forme est:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Est-ce vrai? Si cette matrice était engendré par l'ensemble donné on aurait

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ a & b \end{bmatrix}.$$

qui donnerait $a = 0$ et $b = 0$ (deuxième rangée), tandis qu'on a aussi $a + b = 1$ (première rangée à gauche). C'est une contradiction, donc aucune solution, donc la matrice choisie est valide et on a:

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \right\}.$$

La proposition ci-haut permet de conclure qu'on peut l'ajouter:

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

est indépendant.

---

[3] Si ce n'était pas indépendant dans $\mathbb{R}^2$, on ne pourrait certainement pas trouver un ensemble indépendant en ajoutant des vecteurs!

[4] Coment trouver en tel exemple en général? Si vous voulez en comprendre un peu plus, voir la section “commentaires” au fin

Maintenant on a un "meilleur" ensemble, mais est-ce maximal? Sinon, on pourrait trouver un vecteur qui n'est pas engendré par notre ensemble. On cherche alors un vecteur qui n'est pas combinaison linéaire:

$$a\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + \begin{bmatrix}1 & 0\\1 & 0\end{bmatrix} + c\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} = \begin{bmatrix}a+b+c & 0\\a & b\end{bmatrix}$$

Un exemple d'une matrice qui n'est pas de cette forme est:

$$\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$$

Encore, il y a beaucoup d'autre exemples. Si cette matrice était engendré par les autres on aurait

$$\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} = \begin{bmatrix}a+b+c & 0\\a & b\end{bmatrix}$$

ce qui donne (entre autres) l'équation $1 = 0$ (première rangée à droite). Donc aucune solution, la matrice choisie est valide et on a:

$$\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} \notin \operatorname{span}\left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\1 & 0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\right\}.$$

La proposition ci-haut permet deconclure qu'on peut l'ajouter:

$$\left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\1 & 0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\right\}$$

est indépendant.

Est-ce *maintenant* maximal? Sinon, on pourrait trouver un vecteur qui n'est pas engendré par l'ensemble, en autre mots un vecteur pas de la forme

$$a\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + \begin{bmatrix}1 & 0\\1 & 0\end{bmatrix} + c\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + d\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} = \begin{bmatrix}a+b+c+d & d\\a+d & b+d\end{bmatrix}$$

On a un peu de difficulté à trouver une matrice qui n'est pas de cette forme. Il y a deux raisons possibles: soit on n'a pas cherché assez, soit il n'existe réellement pas une telle matrice. Le deuxième cas équivaut à dire que toute matrice *est* de cette forme, c'est-à-dire que l'ensemble engendre $\mathbb{M}_{2,2}$. Si oui alors on devrait être capable de résoudre

$$\begin{bmatrix}x & y\\z & w\end{bmatrix} = a\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + \begin{bmatrix}1 & 0\\1 & 0\end{bmatrix} + c\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + d\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} = \begin{bmatrix}a+b+c+d & d\\a+d & b+d\end{bmatrix}$$

pour tout x, y, z, w. C'est le système d'équations:

$$x = a+b+c+d \qquad y = d \qquad z = a+d \qquad w = b+d$$

On voit que $d = y$, et donc $a = z - d = z - y$ et $b = w - d = w - y$. Il ne reste que c, qu'on isole de la première équation: $c = x - a - b - d = x - (z-y) - (w-y) - (y) = x+y-z-w$. On a trouvé une solution pour a, b, c, d, donc l'ensemble engendre $\mathbb{M}_{2,2}$. Conclusion: il n'existe aucune autre matrice à ajouter, car il n'existe aucune autre matrice qui n'ext pas engendré par l'ensemble qu'on a déja. Ajouter n'importe quelle autre matrice donnerait un ensemble dépendant. Donc on a un ensemble indépendant maximal pour $\mathbb{M}_{2,2}$:

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + c \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + d \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$$

NB: Comme auparavant, j'ai fait des choix arbitraires. On aurait pu choisir d'autres matrices à ajouter ''a chaque étape, pour donner d'autres ensembles indépendants maximales. Par contre on aurait chaque fois quatre matrices...

C'était un exemple long, mais on a découvert un théorème:

Théorème 7. *Un ensemble indépendant maximal dans V engendre V.*

Ce n'est rien d'autre que la proposition ci-haut dit de façon différent: c'est maximal parce qu'on ne peut plus rien ajouter, parce qu'il n'y a pas de vecteur non-engendré, parce que l'ensemble est engendrant.

Pour résumer, on a l'algorithme suivant.

Algorithme 8. *Soit $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ un ensemble de vecteurs dans V qui est indépendant.*

On peut trouver un ensemble indépendant maximal qui contient cet ensemble en suivant les étapes suivantes.

(0) Vérifier si l'ensemble est indeépendant. Si non, ARRÊTER, la tâche est impossible.

On répète alors les suivants:

(1) Vérifier si l'ensemble engendre V. Si oui, ARRÊTER, l'ensemble est indépendant maximal.
(2) Si l'ensemble n'engendre pas V, on peut trouver un vecteur dans V qui n'est pas engendré par l'ensemble. Ajouter ce vecteur à l'ensemble.
(3) Retourner à l'étape (1).

COMMENTAIRE: COMMENT AUGMENTER

On peut commencer avec un ensemble indépendant pour trouver une base, mais comment est-ce qu'on trouve les nouveaux vecteurs à chaque étape? On cherche un vecteur qui n'est *pas* dans l'espace engendré par l'ensemble qu'on a déja.

Dans l'exemple ci-haut, j'ai "deviné" une matrice pour ensuite vérifier que c'est pas dans l'espace engendré. "Deviner" n'est pas si mal! Géometriquement, on a l'image suivant. On a un espace engendré par deux vecteurs; c'est un peu comme une droite. On cherche un vecteur dans un espace plus grand; c'est un peu comme un plan. Deviner au hasard un point dans un plan, esperant qu'on ne devine pas un point sur une droite...on a une bonne chance. C'est vrai, mais ça ne répond pas vraiment à la question.

Voici peut-être une meilleure réponse. On se rappele qu'on cherchait une matrice qui n'était pas dans

$$\operatorname{span}\left\{\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\right\}$$

Déterminons cet espace engendré précisément!

$$\begin{aligned}
\operatorname{span}\left\{\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\right\} &= \left\{a\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + b\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix} \mid a, b \in \mathbb{R}\right\}\\
&= \left\{\begin{bmatrix}a+b & 0\\ b & a\end{bmatrix} \mid a, b \in \mathbb{R}\right\}\\
&= \left\{\begin{bmatrix}r & s\\ t & u\end{bmatrix} \mid r = a+b, s = 0, t = b, u = a \text{ et } a, b \in \mathbb{R}\right\}
\end{aligned}$$

On a le système d'équations:

$$\begin{aligned}
r &= a + b\\
t &= b\\
u &= a\\
s &= 0
\end{aligned}$$

Ici, on montre r, s, t, u en fonction de a, b: sachant, a, b, on connait r, s, t, u. Mais ce systéme est équivalent (??) à:

$$\begin{aligned}
a &= u\\
b &= t\\
0 &= r - u - t\\
0 &= s
\end{aligned}$$

Ici, on montre a, b en fonction de r, s, t, u: sachant r, s, t, u, on connait a, b. Mais les deux dernières équations paraissent un peu bizarre. Elles disent que, afin d'avoir un a, b, il faudrait respecter certaines conditions:

$$\operatorname{span}\left\{\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}\right\} = \left\{\begin{bmatrix}r & s\\ t & u\end{bmatrix} \mid r - t - u = 0, s = 0\right\}$$

On cherche une matrice qui n'est pas dans l'espace engendré, donc qui ne respecte pas les conditions. Donc une matrice $\begin{bmatrix}r & s\\ t & u\end{bmatrix}$ où soit $r \neq t + u$ ou $s \neq 0$. Ici t et u sont des paramètres. Si on cherche une matrice qui *est* engendré par les deux matrices données, il faut choisir $s = 0$ et $r = t + u$. Si on cherche une matrice qui n'est *pas* engendré par les deux matrices données, il faut choisir soit $s \neq 0$ ou

$r \neq t+u$. C'est exactement ce que j'ai fait pour "deviner" la matrice dans l'exemple: j'ai choisi $t = u = 0$, et j'ai fait exprès pour mettre $r = 1 \neq (0) + (0)$.

Question: Que veut dire "ce système est équivalent…"? La transformation entre les deux systèmes d'équations linéaires est exactement une application la méthode de Gauss-Jordan: on le verra bientôt.

Mike Newman Université d'Ottawa automne 2008

BASES

En cherchant un ensemble engendrant minimal pour un espace vectoriel, on est mené à l'indépendance. En cherchant un ensemble indépendant maximal dans un espace vectoriel on est mené à l'engendrance. Il semble que le cas intéressant serait un ensemble qui est indépendant dans l'espace et qui engendre l'espace.

Définition 1. *Une* base *pour un espace vectoriel V est un ensemble de vecteurs dans V qui est à la fois indépendant dans V et qui engendre V.*

Afin de vérifier si un ensemble est une base, on a alors deux choses à faire:

1) Montrer que l'ensemble est indépendant dans V.
2) Montrer que l'ensemble engendre V.

Exercice: Montrer que $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ est une base pour $\mathbb{R}^2$. C'est-à-dire montrer directement que c'est indépendant dans $\mathbb{R}^2$ et que ça engendre $\mathbb{R}^2$.

On connaît deux autres méthodes de trouver des bases.

Puisqu'un ensemble engendrant minimal pour V est aussi indépendant (un théorème précédent), c'est aussi une base pour V. En revanche, une base pour V engendre V, et il n'existe aucune vecteur dans la base engendré par les autres (pourquoi?), donc c'est un ensemble engendrant minimal pour V. En autres mots, un ensemble engendrant minimal pour V est une base pour V et *vice-versa*.

Puisqu'un ensemble indépendant maximal est aussi engendrant, c'est aussi une base. En revanche, une base pour V est indépendante, et il n'existe aucun vecteur dans V qui n'est pas engendré par la base (pourquoi?), donc une base est un ensemble indépdendant maximal dans V. En autres mots, un ensemble indépendant maximal dans V est une base pour V et *vice-versa*.

Théorème 2. *Soit V un espace vectoriel.*

$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est une base pour V*
$\iff$ $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est indépendant dans V et engendrant pour V*

$\iff \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est engendrant minimal pour* V
$\iff \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est indépendant maximal dans* V

On a trois façons de montrer qu'un ensemble est une base: en choisir celle qui est le plus facile selon la question!

Des exemples du chapitre précédent montre alors que $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ est une base pour $\mathbb{R}^2$ et que $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ est une base pour $\mathbb{M}_{2,2}$. Revoyant ces exemples, on pense que typiquement il y aurait plusieurs bases pour un espace, en fait un nombre infini!

Exercice: Donner un nombre infini de bases pour $\mathbb{R}^2$. Du moins, décrire un tel exemple.

Si tu as completé les exercices précédantes, tu as peut-être constaté qu'une base en $\mathbb{R}^2$ possède toujours 2 vecteurs. Est-ce vrai? Voyons.

Posons que $\{\mathbf{x}\}$ est une base pour $\mathbb{R}^2$. Certainment un ensemble d'un seul vecteur peut être indépendant dans $\mathbb{R}^2$ (e.g., $\mathbf{x} \neq \mathbf{0}$). Mais alors $\{\mathbf{x}\}$ engendrerait une droite passant par l'origine, et n'engendrait *pas* tout l'espace vectoriel $\mathbb{R}^2$. On voit que n'importe quel ensemble de moins de deux vecteurs dans $\mathbb{R}^2$ ne puisse l'engendrer. Donc certainement un ensemble de *moins de deux* vecteurs ne pourrait être une base pour $\mathbb{R}^2$.

Posons que $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ est base de $\mathbb{R}^2$. Certainement un ensemble de trois vecteurs peut engendrer $\mathbb{R}^2$ (exemple?). Mais pourrait cet ensemble être indépendant? Si oui, alors $\{\mathbf{x}, \mathbf{y}\}$ serait indépendant aussi. Mais alors deux vecteurs non-parallèles dans un plan engendreront le plan, qui ferait que $\mathbf{z}$ est engendré par $\{\mathbf{x}, \mathbf{y}\}$, qui ferait que $\{\mathbf{x}, \mathbf{y}\,\mathbf{z}\}$ est dépendant. Donc certainement un ensemble de *plus que deux* vecteurs ne pourraît être une base pour $\mathbb{R}^2$.

La seule possibilité pour une base en $\mathbb{R}^2$: deux vecteurs.

Un argument similaire s'applique en $\mathbb{R}^3$, et on présume (si notre intuition géométrique est valide) pour $\mathbb{R}^n$ aussi.

On constate trois choises:

1) Il semble avoir un minimum pour le nombre de vecteurs qui engendrent un espace
2) Il semble avoir un maximum pour le nombre de vecteurs indépendants dans un espace
3) Il semble que le minimum et le maximum sont égaux

Le résultat suivant aidera à préciser ceci.

Théorème 3. *Soit* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ *un ensemble indépendant dans* V *et* $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\}$ *un ensemble qui engendre* V*.*

Alors $k \leq l$*.*

En mots: la taille d'un ensemble indépendant dans V *est au plus la taille d'un ensemble engendrant pour* V*.*

C'est un résultat fondamental. Pour ceux qui s'intéressent, on donne ici une preuve, mais c'est optionnel. Pour ceux qui ne s'intéressent pas, passez directement au symbole $\Box$.

Preuve: Présumons qu'on a $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ indépendant dans V et $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\}$ engendrant pour B avec $k > l$.

Puisque $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\}$ engendre V il existe des scalaires $a_1, a_2, \ldots, a_l$ tel que

$$\mathbf{x}_1 = a_1\mathbf{y}_1 + a_2\mathbf{y}_2 + \cdots + a_l\mathbf{y}_l$$

Puiqsue $\mathbf{x}_1 \neq \mathbf{0}$ (e.g., que $\{\mathbf{x}_1\}$ est indépendant) il y a au moins un des $a_1, a_2, \ldots, a_l$ qui n'est pas zéro; présumons que c'est a_l. On peut donc récrire cette dépendance comme

$$y_l = \frac{1}{a_l}\left(a_1\mathbf{y}_1 + a_2\mathbf{y}_2 + \cdots + a_{l-1}\mathbf{y}_{l-1} - \mathbf{x}_1\right)$$

Donc $y_l \in \operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{l-1}, \mathbf{x}_1\}$ et alors

$$\operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{l-1}, \mathbf{x}_1\} = \operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_l\} = V.$$

On continue: il existe des scalaires $b_1, b_2, \ldots, b_l$ tel que

$$\mathbf{x}_2 = b_1\mathbf{y}_1 + b_2\mathbf{y}_2 + \cdots + b_{l-1}\mathbf{y}_{l-1} + b_l\mathbf{x}_1$$

Puisque $\{\mathbf{x}_1, \mathbf{x}_2\}$ est indépendant, on sait qu'il y a au moins un des $b_1, b_2, \ldots, b_{l-1}$ qui n'est pas zéro; autrement on aurait une dépendance entre $\{\mathbf{x}_1, \mathbf{x}_2\}$. Présumons que c'est $b_{l-1} \neq 0$. On peut donc récrire cette dépendance comme

$$y_{l-1} = \frac{1}{a_{l-1}}\left(a_1\mathbf{y}_1 + a_2\mathbf{y}_2 + \cdots + a_{l-2}\mathbf{y}_{l-2} - a_l\mathbf{x}_1\right)$$

Donc $y_{l-1} \in \operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{l-2}, \mathbf{x}_1, \mathbf{x}_2\}$ et alors

$$\operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{l-2}, \mathbf{x}_1, \mathbf{x}_2\} = \operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{l-1}, \mathbf{x}_1\} = V.$$

On continue de cette manière, remplacant les $\mathbf{y}$ par les $\mathbf{x}$. Puisque $k > l$ on aurait éventuellement

$$\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l\} = V$$

Donc $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_l\}$ engendre V. C'est indépendant, donc c'est une base pour V et alors (théorème précédent) c'est un ensemble indépendant maximal dans V. C'est une contradiction: on peut ajouter $\{\mathbf{x}_{l+1}, \ldots, \mathbf{x}_k\}$ pour obtenir un ensemble maximal plus grand. Donc on a prouvé que $k \leq l$. $\Box$

Bienvenu à ceux qui n'ont pas lu la preuve, et félicitations à ceux qui l'ont! On continue.

On a plusieurs conséquences de ce théorème.

Exemple: On connaît que $\left\{\begin{bmatrix}1 & 0 & 0\end{bmatrix}^T, \begin{bmatrix}0 & 1 & 0\end{bmatrix}^T, \begin{bmatrix}0 & 0 & 1\end{bmatrix}^T\right\}$ est indépendant dans $\mathbb{R}^3$ (une petite exercice!). Donc n'importe quel ensemble de deux vecteurs ou moins n'engendrent pas $\mathbb{R}^3$.

Exemple: On connaît que $\left\{\begin{bmatrix}1 & 0 & 0\end{bmatrix}^T, \begin{bmatrix}0 & 1 & 0\end{bmatrix}^T, \begin{bmatrix}0 & 0 & 1\end{bmatrix}^T\right\}$ engendre $\mathbb{R}^3$ (une autre mini-exercice!). Donc n'importe quel ensemble de quatre vecteurs ou plus dans $\mathbb{R}^3$ est dépendant.

Exemple: Est-ce que $\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}2\\4\\1\end{bmatrix}, \begin{bmatrix}4\\0\\4\end{bmatrix}, \begin{bmatrix}5\\2\\5\end{bmatrix}\right\}$ est une base pour $\mathbb{R}^3$? On connaît une base de taille trois pour $\mathbb{R}^3$, qui est donc engendrant pour $\mathbb{R}^3$. Notre théorème précédent dit alors que tout ensemble de au moins quatre vecteurs dans $\mathbb{R}^3$ est dépendant, donc n'est certainement pas une base.

Exemple: Est-ce que $\left\{\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\right\}$ est une base pour $\mathbb{M}_{2,2}$? Dans un chapitre précédante, on a trouvé une base de quatre matrices pour $\mathbb{M}_{2,2}$. Étant base, c'est indépendant. Donc aucun ensemble de taille moins que quatre ne pourrait engendrer $\mathbb{M}_{2,2}$. Donc aucun ensemble de taille moins que quatre ne pourrait être base pour $\mathbb{M}_{2,2}$.

Cet exemple suggère un théorème

Théorème 4. *Soit V un espace vectoriel.*

Alors toute base pour V possède le même nombre de vecteurs.

La démonstration est courte: soit $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_s\}$ et $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_t\}$ deux bases. Puisque $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_s\}$ est indépendante et $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_t\}$ engendre V, on a $s \leq t$. Puisque $\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_t\}$ est indépendante et $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_s\}$ engendre V, on a $t \leq s$. Donc $t = s$. De la magie! $\square$

Définition 5. *Soit V un espace vectoriel. Le nombre de vecteurs dans une base pour V est la* dimension *de V; on écrit* $\dim(V)$.

On peut maintenant "augmenter" un théorème:

Théorème 6. *Soit V un espace vectoriel.*

$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est une base pour V*
$\iff$ $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est indépendant dans V et engendrant pour V*
$\iff$ $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est engendrant minimal pour V*
$\iff$ $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est indépendant maximal dans V*
$\iff$ $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est engendrant pour V et* $\dim(V) = n$
$\iff$ $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est indépendant dans V et* $\dim(V) = n$

On a *cinq* façons de montrer qu'un ensemble est une base. Parfois certains sont plus directes que d'autres!

Pourquoi les deux nouvelles méthodes?

Si on connait un ensemble de n vecteurs qui engendrent un espace de dimension n, c'est qu'il existe une base de n vecteurs. Notre ensemble est nécessairement engendrant minimal, sinon on pourrait l'augmenter pour obtenir un ensemble indépendant de taille plus grande que la taille de la base. Donc notre ensemble *est* une base.

Si on connaît un ensemble de n vecteurs indépendants dans un espace de dimension n, c'est qu'il existe une base de n vecteurs. Notre ensemble doit être indépendant maximal, sinon on pourrait la réduire pour obtenir un ensemble engendrant de taille plus petite que la taille de la base. Donc notre ensemble *est* une base.

Or, si on connaît déja la dimension d'un espace vectoriel (par exemple, on a déja déterminé une base), alors afin de trouver une autre base, il suffit de trouver un ensemble soit engendrant ou soit indépendant, mais *ayant le bon nomber de vecteurs.* Dans ce cas on est garanti d'avoir et l'engendrance et l'indépendance en vérifiant qu'une propriété.

Exemple: Montrer que $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ est base pour $\mathbb{R}^2$.

On connaît déja que la dimension est deux, et on voit que ces deux vecteurs sont indépendants dans $\mathbb{R}^2$. Donc c'est une base pour $\mathbb{R}^2$.

Exemple: Montrer que $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\}$ ne contient aucune base pour $\mathbb{R}^3$.

Une base pour $\mathbb{R}^3$ aurait trois vecteurs. Donc le seul espoir est que ces trois vecteurs sont indépendants. Ils ne le sont pas (exercice). Donc ce n'est pas une base pour $\mathbb{R}^3$. Si on enleve des vecteurs, l'ensemble serait trop petit, donc ne contient aucune base pour $\mathbb{R}^3$.

Exemple: Montrer que $\dim(\mathbb{R}^n) = n$. On va construire une base de n vecteurs. On définit $\mathbf{x}_i$ comme le vecteur ayant un 1 dans la i-ième position et des zéros ailleurs.

On vous laisse l'exercice de montrer que c'est indépendant et engendrant (essayer au moins pour une valeur de n: ce serait les systèmes d'équations les plus simples imaginables!). Donc c'est une base. On dit parfois que c'est la *base standard* pour $\mathbb{R}^n$. Voici pour R^2, R^3, R^4:

$$\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$$

Exemple: Montrer que $\dim(\mathbb{M}_{m,n}) = mn$. On construit une base de mn matrices. Pour chaque position, on prend une matrice ayant un 1 dans cette position et des zéros ailleurs. On aurait donc mn matrices au total. On vous laisse l'exercice de montrer que c'est indépendant et engendrant (vraiment, les équations sont super-simples!). Donc c'est une base. On dit parfois que c'est la *base standard* pour $\mathbb{M}_{m,n}$. Voici pour $\mathbb{M}_{3,2}$:

$$\left\{\begin{bmatrix}1&0&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\1&0&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&1&0\end{bmatrix},\begin{bmatrix}0&0&0\\0&0&1\end{bmatrix}\right\}$$

Exemple: Montrer que $\dim(\mathbb{P}_k) = k + 1$. L'exemple cette fois-ci est un ensemble de monômes, un de chaque degré: $\left\{x^k, x^{k-1}, \ldots, x, 1\right\}$. Dit autrement, ce sont les polynômes ayant chacun un coefficient égal à un et les autres égales à zéro (aurait-on vu ceci déja?). Il y en a $k + 1$ car le degré peut être zéro aussi: les constantes. On vous laisse l'exercice de montrer que c'est indépendant et engendrant (non seulement est-ce des équations simplissimo, mais vous les avez déja résolues deux fois). On dit parfois que c'est la *base standard* pour $\mathbb{P}_k$. Voici pour $\mathbb{P}_5$:

$$\left\{1, x, x^2, x^3, x^4, x^5\right\}$$

Théorème 7. *Les espaces* $\mathbb{R}^n$, $\mathbb{M}_{m,n}$ *et* $\mathbb{P}_k$ *possèdent tous des bases.*

Leurs dimensions sont $\dim(\mathbb{R}^n) = n$, $\dim(\mathbb{M}_{m,n}) = mn$ *et* $\dim(\mathbb{P}_k) = k + 1$.

UN CLIN-D'ŒUIL VERS L'INFINI

Exemple: Montrer que $\mathbb{P}$, l'ensemble de touts les polynômes de tout degré, n'a aucune base fini.

C'est un peu différent, mais nos théorèmes nous aident. On connaît que $\{1, x\}$ est un ensemble de polynômes indépenmdant. Mais grâce à un théorème, on sait alors que tout ensemble engendrant possède au moins deux polynômes. Donc toute base (étant un ensemble engendrant) possède au moins deux polynômes. De plus, $\{1, x, x^2\}$ est aussi indépendant, forçant la taille d'une base d'être au moins trois. Même pire, $\left\{1, x, x^2, \ldots, x^k\right\}$ est indépendant pour tout k, forcant une base d'avoir au moins . . . un nombre infini d'éléments. Vu d'une autre manière: n'importe quel ensemble fini de polynômes aurait un degré maximal, et donc les polynômes de degré plus élevé ne seront pas engendré par cet ensemble.

Il n'existe aucune base fini pour $\mathbb{P}$; on dit que $\dim(\mathbb{P}) = \infty$.

Exemple: Montrer que $\mathbb{F}(\mathbb{R})$, l'ensemble de toutes les fonctions réelles, n'a aucune base fini. L'espace $\mathbb{F}(\mathbb{R})$ est énorme: il contient des chose comme $f(x) = \sin(x)$, $g(x) = e^x$, $h(x) = \cos(x^2 - sin(\sqrt{x^2+1}))$, ... Mais il contient aussi les polynômes, et donc contient aussi des ensembles indépendant de toute taille, et donc aucune base fini.

Théorème 8. *Il y a deux types d'espaces vectoriels.*

1) *Ceux ayant des ensembles engendrants fini, et donc des bases fini.*
2) *Ceux n'ayant aucun ensemble engendrant fini, et donc aucune base fini.*

La deuxième sorte dépasse un peu notre cours. Par contre ils sont trés importants. En particulier les propriétés de l'espace $\mathbb{P}$ donnent des méthodes pour approximer numériquement des fonctions beaucoup plus compliqués. T'imagines peut-etre ce que fait ta calculatrice pour évaluer 437×314159: c'est long mais pas compliqué. Par contre ce n'est pas évident comment calculer même $\cos(1)$ par exemple.

SOUS-ESPACES

On a déja vu plusieurs sous-espaces de $\mathbb{R}^n$, $\mathbb{M}_{m,n}$ et $\mathbb{P}_k$, et sans le dire de façon explicite, on a déja trouvé des bases pour plusieurs de ces sous-espaces. Comme exercice, fouiller vos notes!

Exemple: $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + 3z = 0 \right\} = \text{span} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$

On a déja vu cet exemple et on voit que $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$ l'engendre. De plus, ces deux vecteurs sont indépendants (exercice!) et donc forment une base pour W. C'est un sous-espace de $\mathbb{R}^3$ ayant dimension deux.

Exemple: $L = \left\{ t \begin{bmatrix} x \\ 2x \\ -x \end{bmatrix} \mid t \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid y = 2x, z = -x \right\}$.

C'est une droite passant par l'origine en $\mathbb{R}^3$. Une base est $\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \right\}$. C'est un sous-espace de $\mathbb{R}^3$ ayant dimension un. On note que les sous-espaces ont des contraintes, et aussi que le plus de contraintes qu'on a, le plus que la dimension est réduite. Ce serait à revoir.

Pour l'instant, on se contente de la suivante:

Théorème 9. *Soit U un sous-espace de V, et* $\dim(V) = n < \infty$. *Alors*

1) $\dim(U) \leq \dim(V)$

2) $\dim(U) = \dim(V) \iff U = V$

La démonstration est encore courte. Commencer avec n'importe quel ensemble indépendant dans U. Par exemple, un seul vecteur non-zéro, ou même un ensemble vide! Ensuite, on ajoute des vecteurs afin de trouver un ensemble indépendant maximal. Il ne peut exister un ensemble indépendant de plus de n vecteurs dans V, car n est le nomber maximal de vecteurs indépendants dans V. Donc on trouvera un ensemble indépendant maximal dans U (eg une base pour U) ayant au plus n vecteurs.

Si on trouve une base de n vecteurs, c'est necessairemnt un ensemble indépendant maximal dans V aussi, et donc une simultanément une base pour U et pour V. Donc $U = V$ puisqu'ils sont engendrés par le même ensemble. $\square$

Exemple: Soit $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid b + c = 0 \right\}$.

Est-ce que $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ -2 & 3 \end{bmatrix} \right\}$ est une base pour S?

On a déja vu que S est un espace vectoriel (sinon c'est une petite exercice!). Par contre il existe des matrices dans $\mathbb{M}_{2,2}$ qui ne sont pas dans S, par exemple $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (qui ne vérifie pas la condition). Donc $S \neq \mathbb{M}_{2,2}$. On conclut alors que $\dim(S) < \dim(\mathbb{M}_{2,2}) = 4$. Donc la dimension de S est au plus trois. L'ensemble donné, ayant quatre matrices, n'est pas une base pour S.

NB: On n'a pas calculé la dimension de S, on n'en avait pas besoin (exercice: donner une base de trois matrices). On n'a même pas vérifié que les quatre matrices ci-haut sont *dans* S. S'ils ne sont pas, ce n'est certainement pas une base; s'ils sont, ce n'est pas une base.

Exemple: En $\mathbb{R}^3$, la dimension d'une droite passant par l'origine est un et la dimension d'un plan passant par l'origine est deux.

C'est une application de la première partie du théorème. Un plan est sous-espace de $\mathbb{R}^3$ mais n'est pas égal à $\mathbb{R}^3$, donc sa dimension est au plus deux. Tout plan contient une droite passant par l'origine, qui est sous-espace du plan mais n'est pas égal au plan, donc sa dimension est au plus un. Mais toute droite contient un vecteur non-zéro, donc sa dimension est au moins un.

POURQUOI DES BASES? UNICITÉ

C'est beaucoup de théorie, concepts et résultats. On voudrait savoir un peu à quoi ça sert. La réponse est: beaucoup! Premièrement, les vecteurs s'expriment de manière unique par rapport ''a une base.

Théorème 10. *Soit* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *une base pour* V *et* $\mathbf{x} \in V$.

Alors $\mathbf{x}$ *s'écrit de manière unique comme* $\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$.
En autres mots, il existe des scalaires uniques $a_1, a_2, \ldots, a_n$ *tel que* $\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$.

Vous avez déja vu la démonstration comme devoir pour $n = 2$. Puisque c'est une base, c'est engendrant pour V. Donc pour tout $\mathbf{x} \in V$ il existe des scalaires $a_1, a_2, \ldots, a_n$ tel que

$$\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$$

Par contre si on avait

$$\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_n\mathbf{x}_n$$

alors on aurait

$$(a_1 - b_1)\mathbf{x}_1 + (a_2 - b_2)\mathbf{x}_2 + \cdots + (a_n - b_n)\mathbf{x}_n = \mathbf{0}$$

Puisque c'est une base, c'est indépendant, et donc chaque coefficient est nécessairment zéro. Donc $a_1 = b_1$, $a_2 = b_2$, $\ldots a_n = b_n$. $\square$

Définition 11. *Soit* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *une base pour* V *et* $\mathbf{x} \in V$.
Soit $a_1, a_2, \ldots, a_n$ *les scalaires uiques tel que* $\mathbf{x} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$.

Alors on dit que $a_1, a_2, \ldots, a_n$ *sont les* coordonnées de $\mathbf{x}$ par rapport à $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$.

On a déja vu cette idée! Prennons

$$\mathbf{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Alors a, b, c sont exactement les coordonnées de ce vecteur dans $\mathbb{R}^3$ par rapport à la base standard. De plus, prennons

$$p(x) = ax^2 + bx + c = a(x^2) + b(x) + c(1)$$

Alors a, b, c sont exactement les coordonnées de ce vecteur dans $\mathbb{P}_2$ par rapport à la base standard. Les bases permettent de voir que la structure de $\mathbb{R}^3$ et $\mathbb{P}_2$ et pareil: il s'agit de trois coordonnés par rapport à une base.

POURQUOI DES BASES? GÉOMÉTRIE APPLIQUÉ

Cette section dépasse un peu nos connaissances pour le moment, il s'agit de donner une indication générale seulement. On en comprendra plus quand on parle de "transformations linéaires".

Exemple: L'ensemble $\left\{ \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ est une base pour $\mathbb{R}^3$ (exercice). Il s'agit de d'une rotation de la base standard de $30°$ dans le plan xy.

Vous jouez au Frères Mario'nettes sur votre Wee et vous voulez que la scène se tourne de $30°$ pour vous montrer le prochain but. Rien qu'un peu d'elgèbre linéaire: vos coordonnées sont courrament a, b, c, il s'agit de prendre les *mêmes* coordonnées mais *par rapport à la nouvelle base.* La position actuelle est $\mathbf{x}$; la nouvelle est $\mathbf{x}'$.

$$\mathbf{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \rightarrow \quad \mathbf{x}' = a \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} + b \begin{bmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{a\sqrt{3}-b}{2} \\ \frac{a+b\sqrt{3}}{2} \\ c \end{bmatrix}$$

Exemple: L'ensemble $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$ est base pour $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + 3z = 0 \right\}$ (exemple antérieur). C'est un plan incliné, car la dimension est deux. L'ensemble $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ est une base pour $\mathbb{R}^2$.

$\mathbb{R}^2$ c'est votre écran, ou vous avez fait un beau dessin que vous voudriez introduire dans votre nouveu jeu que vous programmez. W c'est un plan incliné sur lequel vous voudriez que ce dessin apparaisse. Un point $\mathbf{x}$ sur votre dessin devrait apparaître où sur W? Au mêmes coordonnées, mais *par rapport à l'autre base.* La position à l'écran est $\mathbf{x}$; la position sur le plan incliné est $\mathbf{x}'$.

$$\mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \rightarrow \quad \mathbf{x}' = a \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2a \\ -3b \\ 1 \end{bmatrix}$$

Malheureusement, personne ne vous a dit qu'une base orthonormale serait préférable. Par contre, la déformation un peu bizarre de votre dessin semble "cool", donc tout n'est pas perdu ...

Il s'agit de transformations linéaires; c'est l'outil principal afin de faire des graphiques en 3D. On verra un peu plus à la fin du cours.

1. Équations Linéaires

Une ÉQUATION LINÉAIRE est une équation de la forme

$$a_1x_1, a_2x_2, \ldots, a_nx_n = b$$

où les valeurs $a_1, a_2, \ldots, a_n$ ainsi que b sont connues et les $x_1, x_2, \ldots, x_n$ sont les variables. On dit typiquement que les a_{ij} sont les COEFFICIENTS du système et les b_i sont les CONSTANTES. On dénotera (typiquement) le nombre de variables par n et le nombre d'équations par m. Une SOLUTION est une liste de valeurs pour chaque variable, tel que la substitution satisfait l'équation. Un SYSTÈME D'ÉQUATIONS est un ensemble de tels équations. En générale, un système est de la forme:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots &= \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Une solution d'un système d'équations est une liste de valeurs pour chaque variable tel que la substitution satisfait *chaque* équation. La SOLUTION GÉNÉRALE d'un système est l'ensemble de toutes les solutions du système.

Ce ne sont pas des concepts nouveaux. Par exemple,

$$3x_1 + 5x_2 - x_3 = 0$$

est une équation ayant $(1, 0, 3)$ et $(2, t, 5t + 6)$ comme solutions (il y en a d'autres). Aussi

$$\begin{aligned} 3x_1 + 5x_2 - x_3 &= 0 \\ 5x_2 - x_3 &= -6 \end{aligned}$$

est un système ayant $(2, t, 5t + 6)$ comme solution mais pas $(1, 0, 3)$. Ici, toutes les solutions sont de la forme $(2, t, 5t+6)$: c'est la solution générale (tentez de prouver ceci! sinon on verra très bientôt.)

Un système est CONSISTANT si il existe (au moins) une solution, et INCONSISTANT si il n'existe aucune solution. Note que dans un système inconsistant *certains* des équations peuvent possèder une solution, mais pas tous!

Un système est HOMOGÈNE si toutes les constantes (c'est-à-dire les b_i ci-haut, pas les a_{ij}) sont zéro. Notre premier résultat est simple mais utile.

Proposition 1. *Un système homogène est toujours consistant.*

La démonstration est directe: un système homogène possède toujours la solution triviale, c'est-à-dire toute variable zéro.[1] Si vous ne comprenez pas ceci, écrire un exemple d'un système homogène et substituer zéro pour chaque variable. Par exemple, on voit que $x = y = z = 0$ est solution pour la suivante:

$$\begin{aligned} x + 7y - 3z &= 0 \\ 2x + y &= 0 \\ y - z &= 0 \\ x + y + z &= 0 \end{aligned}$$

2. Solutions

Exemple: Solutionner

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 2 \end{aligned}$$

On propose que $x = -\frac{3}{2}t$, $y = t+2$, $z = t$ est une solution. Ceci se vérifie facilement en substituant dans chaque équation:

$$\begin{aligned} 2(-\frac{3}{2}t) + 3(t) &= -3t + 3t = 0 \\ (t+2) - (t) &= t + 2 - t = 2 \end{aligned}$$

On montre maintenant que *toutes* les solutions sont de cette forme. La première équation dit que $x = -\frac{3}{2}z$ et la deuxième équation dit que $y = z+2$. Ceci détermine les valeurs pour x et y selon z. On peut alors faire un choix arbitraire de $z = t \in \mathbb{R}$, donnant $x = -\frac{3}{2}t$ et $y = t + 2$. La solution générale est alors

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{3}{2}t \\ t+2 \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} -\frac{3}{2}t \\ t \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{3}{2} \\ 1 \\ 1 \end{bmatrix}$$

C'est en FORME VECTORIELLE; d'habitude on la préfère.

Certaines formes de systèmes sont plus faciles à solutionner que d'autres.

Exemple: Trouver toutes les solutions au système suivant.

$$\begin{aligned} 3x + 2y - 5z &= -2 \\ y + 2z &= 5 \\ 3z &= 15 \end{aligned}$$

[1] Est-ce qu'il y a un lien avec l'indépendance ici?

C'est plus facile de commencer en bas. On voit que obligatoirement $z = 15/3 = 5$. Passant à la deuxième équation, on voit que obligatoirement $y = 5-2z = 5-2(5) = -5$. Passant à la première on voit que obligatoirement $x = (-2-2y+5z)/3 = (-2-2(-5)+5(5))/3 = 11$. La solution unique est

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \\ 5 \end{bmatrix}$$

On observe deux choses: puisque certaines variables n'apparaissent pas dans toutes les équations, on a pu trouver la solution une équation à la fois. Aussi, c'était beaucoup plus facile de commencer en bas. La simplicité provient et de la *forme* du système et de *l'ordre* dans lequel on solutionne. Note que la solution elle-même ne dépend ni de la forme ni de l'ordre: toute méthode (valide!) de solutionner ce système donnerait exactement la même solution. C'est une question d'efficacité: on cherche une méthode aussi générale que possible.

Exemple: Trouver la solution générale du système suivant.

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 2 \\ 0 &= -2 \end{aligned}$$

On commence en bas... et on voit obligatoirement que $0 = -2$. C'est une ÉQUATION CONTRADICTOIRE, ou en vocabulaire plus direct, c'est faux. Peu importe les valeurs de x, z, c'est impossible. Donc aucune solution.

Ce n'est pas vrai que tout système peut se solutionner aussi directement que ces deux exemples. Par contre, on verra que dans tout système se cache un système *équivalent* (?) qui peut se solutionner directement.

Il y aura alors deux étapes pour solutionner un système. Premièrement la transformer en forme équivalent mais plus simple; deuxièmement, trouver la solution directement en utilisant cette forme plus simple.

3. Matrices

Avant de continuer, on fera une simplification de notation. Un système linéaire est complètment déterminé par ces coefficients et ces constantes. La seule raison qu'on doit écire chaque variable ("x", "y", etc) c'est pour ne pas les confondre. La MATRICE DES COEFFICIENTS d'un système est une matrice ayant une rangée pour chaque équation et une colonne pour chaque variable; l'élément de la matrice dans la position i, j est le cofficient de la j-ième variable dans la i-ième équation. En voici un exemple d'un système et sa matrice des coefficients.

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 2 \\ -2x + 3y - 6z &= 0 \end{aligned} \qquad \begin{bmatrix} 2 & 0 & 3 \\ 0 & 1 & -1 \\ -2 & 3 & -6 \end{bmatrix}$$

Les zéros correspondent aux variables qui ne sont *pas* présentes dans une équation. Pour un système de m équations en n variables, la matrice des coefficients serait de taille $m \times n$.

La MATRICE AUGMENTÉE est la matrice des coefficients avec une colonne de plus, correspondant aux constantes; on met une barre verticale pour séparer les coefficients des constantes.[2]

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 2 \\ -2x + 3y - 6z &= 0 \end{aligned} \qquad \left[\begin{array}{ccc|c} 2 & 0 & 3 & 0 \\ 0 & 1 & -1 & 2 \\ -2 & 3 & -6 & 0 \end{array}\right]$$

Pour un système de m équations en n variables, la matrice augmentée serait de taille $m \times (n+1)$. On peut lire une matrice augmentée comme un système d'équations: par exemple la deuxième rangée dit exactement "$0x + 1y - 1z = 2$". On manipule une matrice augmentée comme des équations. Par exemple, on peut additionner la première équation à la dernière, remplaçant ainsi la dernière:

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 2 \\ 3y - 3z &= 0 \end{aligned} \qquad R_3 \to R_3 + R_1 \qquad \left[\begin{array}{ccc|c} 2 & 0 & 3 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & 3 & -3 & 0 \end{array}\right]$$

Ensuite on peut multiplier la dernière par $\frac{1}{3}$:

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 2 \\ y - z &= 0 \end{aligned} \qquad R_3 \to \tfrac{1}{3} R_3 \qquad \left[\begin{array}{ccc|c} 2 & 0 & 3 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & 1 & -1 & 0 \end{array}\right]$$

Ensuite on peut soustraire la deuxième de la dernière:

[2] Le manuel ne met aucune barre, ce qui cause parfois une confusion. En classe, devoirs, tests, examens, etc, on la mettrais toujours.

$$\begin{aligned} 2x + 3z &= 0 \\ y - z &= 2 \\ 0 &= -2 \end{aligned} \qquad R_3 \to R_3 - R_2 \qquad \left[\begin{array}{ccc|c} 2 & 0 & 3 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & -2 \end{array}\right]$$

En additionnant des équations (rangées) et multipliant des équations (rangées) par des constantes, on n'a pas changé la solution générale (pourquoi?). Mais la solution inclut maintenant "$0 = -2$", une impossibiolité, donc il n'y a aucune solution. On a pu trouver un système "simple" caché dans un système "compliqué". C'est exactement l'idée générale qu'on cherche

On donnera dorénavant la matrice augmentée au lieu du système; c'est une forme plus compacte. Mais garder l'équivalence en tête!

Terminologie que vous devriez comprendre: équation linéaire, système linéaire, solution, consistant, inconsistant, système homogène, matrice des coefficients, matrice augmentée

1. Forme Échelonnée

On a déja vu un exemple d'un système "simple"; la revoici, en matrice augmentée

$$\left[\begin{array}{ccc|c} 3 & 2 & -5 & -2 \\ 0 & 1 & 2 & 5 \\ 0 & 0 & 3 & 15 \end{array}\right]$$

C'était facile à solutionner puisqu'on pouvait procéder une équation à la fois, résolvant pour z, y, x en ordre. La variable pour laquelle on à résolue correspondait chaque fois au premier non-nul dans chaque rangée. Ce sont les PIVOTS. Dans cette matrice, les pivots sont "échelonnés" de gauche à droite, ce qui veut dire que quand on résout pour z, il n'y a aucun x et y; quand on résout pour y il n'y a aucun x. Formellement:

Définition 1. *Une matrice est en* FORME ÉCHELONNÉE *si*

(1) Toute rangée nulle se trouve en bas (s'il y en a)
(2) Tout pivot est égal à 1
(3) Tout pivot se trouve à la droite des pivots supérieurs.

On voit que la matrice précédante est presque en forme échelonnée: deux des pivots ne sont pas 1. On peut la transformer en forme échelonnée en multipliant la première rangée par $\frac{1}{3}$ et la troisième par $\frac{1}{3}$. Ceci ne change pas la validité des équations correspondantes (pourquoi?). Ici, les pivots sont en $\boxed{\text{boîte}}$.

$$\left[\begin{array}{ccc|c} \boxed{1} & \frac{2}{3} & -\frac{5}{3} & -\frac{2}{3} \\ 0 & \boxed{1} & 2 & 5 \\ 0 & 0 & \boxed{1} & 5 \end{array}\right] \qquad R_1 \to \frac{1}{3}R_1, \quad R_3 \to \frac{1}{3}R_3$$

Exercice: Les matrices suivantes ne sont *pas* en forme échelonnée. Pourquoi? Comment est-ce qu'on pourrait les transformer en forme échelonnée sans changer la validité des équations? Écrire le système d'équations correspondant à chaque matrice avant et après votre transformation. Expliquer pourquoi ils sont équivalents.

Que veut dire "équivalent"? Aussi, identifier les pivots avant et après votre transformation.

$$\left[\begin{array}{cccc|c} 0 & 1 & -7 & 5 & 2 \\ 1 & 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 1 & 8 \end{array}\right] \qquad \left[\begin{array}{cccc|c} 1 & 0 & -7 & 5 & 2 \\ 1 & 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 1 & 8 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 0 & -7 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 8 \end{array}\right]$$

2. Solutionner

Exemple: Donner la solution générale.

$$\left[\begin{array}{cccc|c} \boxed{1} & 3 & -7 & 4 & 2 \\ 0 & 0 & \boxed{1} & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

La matrice augmentée est en forme échelonnée car chaque pivot est 1, et chaque pivot est à la droite des pivots supérieurs. Aussi, il y a deux rangées nulles, mais elles sont en bas. Elles correspondent à l'équation "$0 = 0$", qu'on peut ignorer.[1] Il y a deux pivots, en $\boxed{\text{boîte}}$. Donc on pourra résoudre pour deux variables. Chaque colonne *avec* pivot correspond à une variable liée. Il y a quatre variables au total (les colonnes de la matrice des coefficients). Chaque colonne *sans* pivot correspond à une variable libre, c'est-à-dire un paramètre. On pourra résoudre pour les variables liées en terms des paramètres. On met $x_2 = s$ et $x_4 = t$ pour les paramètres, et on lit chaque rangée avec pivot, de bas en haut afin de trouver les variables liées. Donc

$$\begin{aligned} x_3 + 5x_4 = 3 \quad &\rightarrow \quad x_3 = 3 - 5x_4 \\ &\qquad\;\;\; = 3 - 5(t) \\ x_1 + 3x_2 - 7x_3 + 4x_4 = 2 \quad &\rightarrow \quad x_1 = 2 - 3x_2 + 7x_3 - 4x_4 \\ &\qquad\;\;\; = 2 - 3(s) + 7(3 - 5t) - 4(t) \\ &\qquad\;\;\; = 23 - 3s - 39t \end{aligned}$$

ce qui donne comme solution générale:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 23 - 3s - 39t \\ s \\ 3 - 5t \\ t \end{bmatrix} = \begin{bmatrix} 23 \\ 0 \\ 3 \\ 0 \end{bmatrix} + s \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -39 \\ 0 \\ -5 \\ 1 \end{bmatrix}$$

Avec un peu de pratique on peut lire la solution générale directment de la matrice sans devoir écire les équations. Mais si vous n'êtes pas encore aussi confortable, n'hésitez pas à écrire les équations!

[1] Pourquoi alors est-ce qu'on écrit des rangées nulles, si ce n'est que pour les ignorer? En réailté, une rangée nulle est souvent caché dans un système, un peu comme l'équation "$0 = -2$" était caché dans un système du chapitre précédent. Ce n'est typiquement qu'après avoir transformé le système qu'on les découvres, et à ce moment on peut les ignorer.

Comment lire la solution d'une matrice augmenté en forme échelonée:

(1) Est-ce qu'il y a une rangée contradictoire? Si oui, ARRÊTER. Aucune solution.
(2) Regarder dans les rangées afin de trouver les pivots (premiers non-nuls).
(3) Regarder dans les colonnes pour trouver les variables liées (colonnes avec pivots) et les variables libres (colonnes sans pivots).
(4) Poser chaque variable libre comme paramètre.
(5) Résoudre pour chaque variable liée, de bas en haut, substituant au besoin les autres variables liées qu'on a déja déterminées.
(6) Donner la réponse finale, pour *toutes* les variables.

Chaque fois qu'on résoud pour une variable liée, ce serait en termes de paramètres et d'autres variables liées qu'on connaît déja (c'eat pour ceci qu'on commence en bas!) — par exemple, le x_1 était en termes des deux paramètres ainsi que la variable x_3 qu'on connaissait déja.

Qu'est ce qu c'est une rangée contradictoire? Une rangée qui correspond à une équation qui est *toujours fausse*, par exemple $0 = -2$. Dans la matrice en forme échelonnée, ce serait une rangée du genre $\left[\begin{array}{cccc|c} 0 & 0 & \dots & 0 & K \end{array}\right]$ où $K \neq 0$.

Note qu'une rangée nulle, $\left[\begin{array}{cccc|c} 0 & 0 & \dots & 0 & 0 \end{array}\right]$ ne signifie rien, tandis qu'une rangée contradictoire signifie aucune solution. Très différent!

Exercice: Donner la solution générale pour les matrices suivantes:

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right] \qquad \left[\begin{array}{cccc|c} 1 & 3 & -7 & 4 & 2 \\ 0 & 1 & 1 & 5 & 3 \end{array}\right]$$

3. Décrire la Solution Générale

Si on connaît les pivots, on peut décrire la solution avant même de résoudre.

Premièrement, s'il y a une rangée contradictoire, c'est qu'il y a un pivot qui est à la droite de la barre (un? pourquoi pas deux?). Autrement dit, il un pivot de plus dans la matrice augmentée que dans la matrice des coefficients. Par contre si tous les pivots sont à la gauche de la barre, il n'y a aucune rangée contradictoire et il y aura une solution.

Deuxièmement, s'il y a un pivot dans chaque colonne de la matrice des coefficients, il n'y aura aucun paramètre. On pourra donc résoudre pour chaque variable liée, sans choix, et il y aura alors une solution unique. S'il y a une colonne sans pivot (ou plus qu'une!), il y aura (au moins) un paramètre. Donc, dans la solution, il y aura un choix libre parmi $\mathbb{R}$, donc une infinité de solutions.

On connaît le genre de solution en fonction des pivots:

Théorème 2. *Un système d'équations linéaires possède soit aucune solution, une solution unique, ou une infinité de solutions.*

On peut distinguer les trois cas sachant les pivots dans une forme échelonnée. S'il y a un pivot dans la colonne des constantes, il n'y a aucune solution. S'il n'y a aucun pivot dans la colonne des constantes et un pivot dans chaque colonne de la matrice des coefficients, il y aura une solution unique. S'il n'y a aucun pivot dans la colonne des constantes et il manque un pivot dans une colonne de la matrice des coefficients, il y aura une infinité de solutions.

NB: On regarde dans les *rangées* pour savoir *ou* sont les pivots; on regarde dans les *colonnes* pour comprendre leurs signifiance.

Exercice: Identifier les pivots, donner le nombre de variables liées et paramèmetres, et donner le nombre de solutions. Ne pas solutionner!

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 0 \end{array}\right] \qquad \left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & 8 \\ 0 & 0 & 9 \end{array}\right]$$

4. Forme Échelonnée Réduite

La forme échelonnée sert très bien à résoudre un système, et a décrire la solution. Mais on est un peu paresseux: on ne veut pas toujours resubstituer les variables liées qu'on a déja trouvé (par exemple, la substitution de x_3 afin de trouver x_1 ci-haut). Pour un exemple plus grand, eg 50000 équations et variables, ce serait une étape longue, et aussi répetée (il faudrait peut-être substituer x_{50000} dans x_{49999}, x_{49998}, x_{49997} ...). Il y a un autre probleme. Dépendant dans quel ordre on fait les opérations, on pourrait obtenir des formes échelonnées différentes!

Començons avec $x + 2y - z = 3, y + z = 2$. On pourra soit l'écrire comme matrice augmentée toute suite (à gauche) ou soit additionner la deuxième équation à la première (à droite):

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & 1 & 1 & 2 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 3 & 0 & 5 \\ 0 & 1 & 1 & 2 \end{array}\right]$$

Laquelle est "la" forme échelonnée? Réponse: les deux, ainsi qu'une infinité d'autres. On aura donc de la difficulté à savoir quand deux systèmes sont équivalents (eg, ont la même solution générale).

On découvre ainsi la forme échelonnée réduite:

Définition 3. *Une matrice est en* FORME ÉCHELONNÉE RÉDUITE *si*

(1) Toute rangée nulle se trouve en bas (s'il y en a)
(2) Tout pivot est égal à 1
(3) Tout pivot se trouve à la droite des pivots supérieurs.
(4) Tout pivot et le seul non-nul dans sa colonne.

Il y a deux avantages. En résolvant, on trouve chaque variable liée en termes de paramètres *seulement*. On n'a jamais besoin de substituer des variables, et donc ce n'est plus important de commencer en bas! Deuxièmement, c'est unique, dans un sens qu'on précisera bientôt.

Exercice: Lesquels des suivants sont en forme échelonnée réduite? Pourquoi? Sinon, tenter de les transformer en forme équivalent échelonné réduite (encore ce mot "équivalent"! voulant dire quoi?).

$$\left[\begin{array}{cccc|c} 0 & 1 & -7 & 0 & 2 \\ 1 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 8 \end{array}\right] \qquad \left[\begin{array}{cccc|c} 1 & 0 & 8 & -5 & 12 \\ 0 & 1 & 1 & -3 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \qquad \left[\begin{array}{cccc|c} 1 & 0 & -7 & 0 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 8 \end{array}\right]$$

5. Solutionner

Exemple: Donner la solution générale

$$\left[\begin{array}{ccc|c} 1 & 0 & -5 & 6 \\ 0 & 1 & 1 & 3 \end{array}\right]$$

Elle est en forme échelonnée car chaque pivot est 1, et chaque pivot est à la droite des pivots supérieurs. Elle est en forme échelonnée réduite car de plus chaque pivot est le seul non-nul dans sa colonne.

Il y a deux pivots, donc deux variables liées. Une colonne dans la matrice des coefficients ne possède pas de pivot, donc on aura un paramètre: $x_3 = t$. On résoud maintenant pour les variables liées. La première rangée donne $x_1 = 6 + 5t$ et la deuxième rangée donne $x_2 = 3 - t$. Solution générale, en forme vectorielle:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6+5t \\ 3-t \\ t \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 0 \end{bmatrix} + t \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix}$$

C'est l'équation d'une droite dans $\mathbb{R}^3$!

Exemple: Donner la solution générale pour le système suivant:

$$\left[\begin{array}{cccc|c} 1 & 2 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

On ignore la rangée nulle. Il n'y a aucune rangée contradictoire, donc il y a une solution. Il y a deux paramètres, $x_2 = s$ et $x_3 = t$. La première rangée donne que $x_1 = 4 - 2x_2 + 3x_3 = 4 - 2s + 3t$ et la deuxième donne $x_4 = -5$. Solution générale (comme d'habitude, en forme vectorielle):

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 - 2s + 3t \\ s \\ t \\ -5 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ -5 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

C'est l'équation d'un plan en $\mathbb{R}^4$!

On a déja vu qu'on a soit aucune solution, soit une solution unique, soit une infinité de solutions. On peut maintenant en dire plus: la solution générale est soit vide (aucune solution), un point (solution unique), une droite (infinité de solutions), un plan (infinité de solutions), un hyper-plan (infinité de solutions), ... On peut maintenant préciser que les "infinités de solutions" peuvent se classifier selon leur "dimension" (attention! ce n'est pas la dimension d'un espace vectoriel, car une droite n'est pas toujours un espace vectoriel!).

Comment lire la solution d'une matrice augmenté en forme échelonée réduite:

(1) Est-ce qu'il y a une rangée contradictoire? Si oui, ARRÊTER. Aucune solution.
(2) Regarder dans les rangées afin de trouver les pivots (premiers non-nuls).
(3) Regarder dans les colonnes pour trouver les variables liées (colonnes avec pivots) et les variables libres (colonnes sans pivots).
(4) Poser chaque variable libre comme paramètre.
(5) Résoudre pour chaque variable liée, dans n'importe quel ordre.
(6) Donner la réponse finale, pour *toutes* les variables.

On note que pour une matrice en forme échelonnée réduite, il n'y a aucune simplification afin de donner la solution générale. Chaque équation (=rangée) non-nulle possède exactement une variable liée; en résolvant, on met tout de l'autre côté. Ceci veut dire qu'on a une équivalence.

Théorème 4. *Deux matrices en forme échelonnée réduite ont la même solution générale si et seulement si elles sont identiques, sauf peut-être pour des rangées nulles.*

Exercice: Donner la solution générale pour les suivants:

$$\left[\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 7 \\ 0 & 0 & 1 & 2 & -5 \end{array}\right] \qquad \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & 1 & -6 \\ 0 & 1 & 0 & -2 & -8 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Exercice: Décrire la forme échelonnée réduite qui correspond à un système ayant une solution unique. Attention, ce n'est vrai qu'il s'agit toujours de n variables et n éqations!

Terminologie que vous devriez comprendre: pivot, paramètre, variable libre, variable liée, forme échelonnée, forme échelonnée réduite

6. Un Peu Plus Avancé

Pour ceux qui veulent comprendre un peu plus, cette section considère un peu plus ce que c'est une “bonne forme”. C'est une section optionnelle.

Exemple: Donner la matrice augmentée qui correspond aux solutions générales suivantes:

$$L_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 7 \end{bmatrix} + s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \qquad L_2 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 7 \end{bmatrix} + s \begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix}$$

La première correspond aux équations $x_1 + 2s = 5$, $x_2 = s$ et $x_3 = 7$. C'est clair que x_2 est le paramètre, donnant:

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 5 \\ 0 & 0 & 1 & 7 \end{array}\right]$$

La deuxième correspond aux équations $x_1 - 4s = 3$, $x_2 + 2s = 1$ et $x_3 = 7$. Mais maintenant on voit difficilement qui est paramètre!

En réalité $L_1 = L_2$. Ce sont deux formes différents pour la même droite. On a déja remarqué qu'une droite possède plusieurs équations, toutes équivalentes (choisir un point de départ différent mais encore sur la droite ou choisir un vecteur directeur différent mais dans la même direction). Parmi toutes ces formes, il y a une, et exactement une, qui correspondra au système échelonnée réduite.

Comment trouver *cette* équation de la droite? Premièrement il faudrait changer le point de départ en ajoutant un multiple du vecteur directeur pour que la dernière position non-nulle dans ce vecteur annule la position correspondante dans le vecteur du point:

$$\begin{bmatrix} 3 \\ 1 \\ 7 \end{bmatrix} \to \begin{bmatrix} 3 \\ 1 \\ 7 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 7 \end{bmatrix}$$

Ensuite il faudrait multiplier le vecteur directeur par une constante afin que cette position devient 1:

$$\begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix} \to -\frac{1}{2}\begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

Au cas de deux vecteurs directeurs, il faudrait premièrement utiliser un vecteur pour créer un zéro dans l'autre…

On vous invites d'essayer de décrire une méthode générale! C'est une version de la méthode de Gauss-Jordan; revenez à cette question après le prochain chapitre.

Mike Newman Université d'Ottawa automne 2008

1. Équivalence de Systèmes

On connaît comment résoudre des systèmes s'ils sont déja dans une forme "facile", soit échelonnée ou même échelonnée réduite. Mais en générale comment résoudre un système arbitraire, sans aucune forme particulière? On cherche une méthode à transformer un système arbitraire à un système en forme échelonnée réduite.

Définition 1. *Deux systèmes sont équivalents si ils ont exactement la même solution générale.*

Notre transformation devrait toujours nous ammener à un système équivalent. On veut savoir quelles opérations sur un système d'équations vont conserver la solution générale.

Théorème 2. *Soit $E_1, E_2, \ldots, E_m$ un système de m équations en n variables.*

Soit $S = \{\mathbf{u} \in \mathbb{R}^n \mid \mathbf{x}$ est solution de chaque $E_1, E_2, \ldots, E_m\}$ la solution générale.

Alors S est aussi la solution générale pour le système $E_1 + E_2, E_2, E_3, \ldots, E_m$ ainsi que pour le système $tE_1, E_2, \ldots, E_m$ pour tout $t \neq 0$.

On comprend ce résultat comme suit. Qu'un vecteur $\mathbf{u} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}^T$ soit dans S veut dire que les valeurs $u_1, u_2, \ldots, u_n$ vont satisfaire chacune des équations $E_1, E_2, \ldots, E_m$. Alors il vont satisfaire une somme de deux équations aussi, car la somme de deux égalités produit une égalité. Aussi, il vont satisfaire un multiple d'une équation, car un multiple d'une égalité produit aussi une égailté. Pourquoi alors $t \neq 0$? Car une équation *fausse* devient *vraie* lorsqu'on multiplie chaque côté par zéro ($0 = -2$ devient $0 = 0$). Multiplier une équation par zéro équivaut à enlever une contrainte, ce qui change la solution générale!

L'étudiant motivé pourra demander pourquoi est-ce que ce sont exactement les deux opérations d'un espace vectoriel qui sont identifiées dans ce théorème?

2. Réduction à la Forme Échelonnée

On trouve les trois opérations suivantes utiles. On les présente ici comme opérations de rangée, mais on comprend que c'est la même chose que des opérations d'équations.

On a trois opérations de rangée fondamentales.

I.	Échanger deux rangées.	$R_i \rightleftharpoons R_j$
II.	Multiplier une équation par $t \neq 0$.	$R_i \to tR_i$
III.	Additionner un multiple d'une rangée à une autre.	$R_i \to R_i + tR_j$

Ces opérations préservent la solution générale d'un système, car ce sont des multiplications scalaires et additions. Donc si on utilise ces opérations, tout système qu'on obtiendra sera équivalent.

On peut se servir de ces trois opérations, dans n'importe quel ordre, afin d'éventuellement obtenir une forme échelonnée réduite, d'où on pourra lire la solution générale. Mais comment fait-on pour éviter de tourner en rond? On a besoin non seulement de savoir *quelles* opérations sont permises, mais *comment s'en servir*: on a besoin d'un algorithme.

Algorithme 3 (Gauss)**.** *On peut transformer tout système linéaire à une forme équivalente échelonnée comme suit:*

(1) Est-ce que la matrice est nulle? Si oui, ARRÊTER.
(2) Trouver la colonne non-nulle la plus à gauche.
(3) Choisir un élément non-nul dans cette colonne. Utiliser les opérations de type I et II afin de mettre un 1 *dans la première position de cette colonne.*
(4) Utiliser cet élément comme pivot, et à l'aide des opérations de type III annuler tout non-nul en bas *du pivot.*
(5) Ignorer la rangée contenant ce pivot et retourner à l'étape (1)

A chaque répétition on passe à une nouvelle rangée, donc il y a au plus m répétitions. Note qu'à l'étape (4) on s'en sert du pivot pour annuler en bas seulement... c'est une question d'efficacité, on y reviendra dans quelques instants!

Exemple: Donner une forme équivalente échelonnée de la suivante:

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 6 \\ -1 & -1 & 0 & 0 & 4 \\ 2 & 5 & -3 & 1 & -1 \end{array}\right]$$

On suit l'algorithme de Gauss, balayant les colonnes de gauche à droite.

On a déja un 1 dans la première position de la première colonne non-nulle, donc on passe à l'étape (3). Afin d'annuler, on note que l'opération est déterminée: on veut annuler le "-1" donc il faut faire $R_2 \to R_2 + R_1$. En conséquence, il faut fair exactement ceci pour toute la deuxième rangée. Aussi, on veut annuler le "2", donc il faut faire $R_3 \to R_3 - 2R_1$:

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 6 \\ 0 & 2 & -2 & 2 & 10 \\ 0 & -1 & 1 & -3 & -13 \end{array}\right] \quad R_2 \to R_2 + R_1, \quad R_3 \to R_3 - 2R_1$$

On s'organise pour que le prochain pivot devient 1, l'étape (3):

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 6 \\ 0 & 1 & -1 & 1 & 5 \\ 0 & -1 & 1 & -3 & -13 \end{array}\right] \quad R_2 \to \frac{1}{2}R_2$$

On annule en bas avec le deuxième pivot:

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 6 \\ 0 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & -2 & -8 \end{array}\right] \quad R_3 \to R_3 + R_2$$

On multiplie pour que le dernier pivot soit 1:

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 6 \\ 0 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right] \quad R_3 \to -\frac{1}{2}R_3$$

Il n'y a plus de rangées, donc on a fini! La matrice est maintenant en forme échelonnée, et de plus la solution générale n'a pas changée.

On voit qu'il y aura une solution, car il n'y a aucune rangée contradictoire. Il y aura une infinité de solutions, car il y a un paramètre: la troisième colonne n'a pas de pivot.

Exemple: Donner une forme équivalente échelonné de la suivante:

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 2 & 4 & -2 & 2 \\ 3 & 2 & -7 & 3 \end{array}\right]$$

On commence avec un échange de rangées afin de mettre un 1 en haut à gauche:

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 3 & 2 & -7 & 3 \end{array}\right] \quad R_2 \rightleftharpoons R_1, \quad R_1 \to \frac{1}{2}R_1$$

Ensuite on annule en bas:

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & -4 & -4 & 0 \end{array}\right] \quad R_3 \to R_3 - 3R_1$$

Le deuxième pivot et déja en position; on annule en bas:

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 4 \end{array}\right] \qquad R_3 \to R_3 + 4R_2$$

On n'a pas fini: le dernier pivot n'est pas 1. Par contre, c'est dans la colonne des constantes; autrement dit, la dernière rangée dit "$0 = 4$", une contradiction. Il n'y aura aucune solution.

Exercice: Donner une forme équivalente échelonné de la suivante:

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 2 & 4 & -2 & 2 \\ 3 & 2 & -7 & -1 \end{array}\right]$$

Donner le nombre de pivots, variables liées, paramètres, et nombre de solutions.

3. Réduction à la Forme Échelonnée Réduite

On cherchait un algorithme pour aboutir à la forme échelonnée réduite. Ayant déja une méthode pour arriver à forme échelonnée, ce n'est pas long.

Algorithme 4 (Gauss-Jordan)**.** *On peut transformer tout système linéaire à une forme équivalente échelonnée réduite comme suit:*

(0) Transformer en forme échelonnée suivant l'algorithme de Gauss ci-haut.
(1) Trouver le pivot le plus à droite
(2) Utiliser ce pivot pour annuler en haut.
(3) Ignorer la colonne contenant ce pivot et retourner à l'étape (1)

Il s'agit d'annuler "en haut", exactement ce qu'on a évité de faire à l'étape (4) de la méthode de Gauss, mais cette fois de droite à gauche. On comprendra mieux pourquoi on préfère cet ordre dans un exemple.

Exemple: Donner une forme équivalente échelonné réduite de la suivante:

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 6 \\ -1 & -1 & 0 & 0 & 4 \\ 2 & 5 & -3 & 1 & -1 \end{array}\right]$$

On connaît déja une forme échelonnée équivalente:

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 2 & 6 \\ 0 & 1 & -1 & 1 & 5 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

On utilse le pivot dans la quatrième colonne pour annuler en haut:

$$\left[\begin{array}{cccc|c} 1 & 3 & -2 & 0 & -2 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right] \quad R_2 \to R_2 - R_3, \quad R_1 \to R_1 - 2R_3$$

On utilse le pivot dans la deuxième colonne pour annuler en haut:

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & -5 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right] \quad R_1 \to R_1 - 3R_2$$

Note que dans cetter dernière opération on n'a rien dû calculer pour la quatrième colonne, car il s'agit de $0 \to (0) - 3(0)$. On était certain d'avoir deux zéros ici, à cause qu'on a déja annulé en haut dans la quatrième colonne. C'est une opération qu'on a épargné en annulant en haut *maintenant* au lieu qu'à l'étape (4) de la méthode de Gauss. Ce n'est qu'une opération, mais avec une matrice plus grand, ce serait plus!

Exemple: Donner une forme équivalente échelonné de la suivante:

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 2 & 4 & -2 & 2 \\ 3 & 2 & -7 & -1 \end{array}\right]$$

Ce qui ont déja fait l'exercice savent qu'on a une forme échelonnée:

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Il ne reste qu'à annuler en haut avec le pivot dans la deuxième colonne:

$$\left[\begin{array}{ccc|c} 1 & 0 & -3 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad R_1 \to R_1 - 2R_2$$

4. Solutionner

On connaît déja que deux matrices en forme échelonnée réduite on la même solution si et seulement si elles sont identiques. On se souvient aussi que deux matrices augmentées sont équivalents exactement lorsqu'elles ont la même solution générale. Donc, pour chaque matrice il existe exactement une matrice équivalente en forme échelonnée réduite.

Théorème 5. *Toute matrice augmentée possède une et seulement une matrice équivalente en forme échelonnée réduite.*

En conséquence, la définition suivante devient possible.

Définition 6. *Le rang d'une matrice est le nombre de pivots dans sa forme échelonnée réduite. On dit* $\operatorname{rang}(M)$ *pour le rang d'une matrice* M.

En fait, il suffit de trouver une forme échelonnée, car la transformation de échelonnée en échelonnée réduite ne change pas le nombre de pivots.

Théorème 7. *Soit* M *une matrice. N'importe quelle matrice équivalente à* M *en forme échelonnée possède* $\operatorname{rang}(M)$ *pivots.*

On a déja caractérisé le nombre de solutions d'une matrice augmentée en termes des pivots. (On ne l'a pas dit, mais à ce moment on avait déja prouvé le théorème précédent.) On peut maintenant le faire en terme du rang. Soit A la matrice des coefficients d'un système, et $A|b$ la matrice augmentée. L'absence d'une rangée contradictoire veut dire que touts les pivots sont à la gauche de la barre, c'est-à-dire que $\operatorname{rang}(A|b) = \operatorname{rang}(A)+1$. La présence d'une rangée contradictoire veut dire qu'il y a un pivot dans la colonne des constantes (mais pas plus qu'un; pourquoi?), et donc $\operatorname{rang}(A|b) = \operatorname{rang}(A)$. L'absence d'un paramètre veut dire que chaque colonne de la matrice des coefficents possède un pivot, donc $\operatorname{rang}(A) = n$. La présence d'un paramètre veut dire qu'il y a une colonne dans la forme échelonnée réduite de A qui ne possède pas un pivot, donc $\operatorname{rang}(A) < n$. Tout ceci permet de caractériser le nombre de solutions comme suit:

Théorème 8. *Soit* $[A|b]$ *la matrice augmentée d'un système linéaire. Alors*

(1) Le système possède aucune solution $\iff \operatorname{rang}(A|b) = \operatorname{rang}(A) + 1$
(2) Le système possède une solution unique $\iff \operatorname{rang}(A|b) = \operatorname{rang}(A) = n$
(3) Le système possède infinité de solutions $\iff \operatorname{rang}(A|b) = \operatorname{rang}(A) < n$

De plus on pourrait dire que la "dimension" de la solution générale est $n - \operatorname{rang}(A)$.

Pour finir, un dernier exemple ayant un paramètre dans la matrice des coefficients.

Exemple: Pour quelles valeurs de k est-ce que le suivant possède aucune solution, une solution unique, une infinité de solutions?

$$\left[\begin{array}{ccc|c} 1 & 2 & k & 1 \\ 1 & k & -1 & -2 \\ 1 & 1 & -1 & 4 \end{array}\right]$$

On réduit:

$$\left[\begin{array}{ccc|c} 1 & 2 & k & 1 \\ 0 & k-2 & -1-k & -3 \\ 0 & -1 & -1-k & 3 \end{array}\right] \quad R_2 \to R_2 - R_1, \quad R_3 \to R_3 - R_1$$

$$\left[\begin{array}{ccc|c} 1 & 2 & k & 1 \\ 0 & -1 & -1-k & 3 \\ 0 & k-2 & -1-k & -3 \end{array}\right] \quad R_2 \rightleftharpoons R_3$$

$$\left[\begin{array}{ccc|c} 1 & 2 & k & 1 \\ 0 & 1 & 1+k & -3 \\ 0 & k-2 & -1-k & -3 \end{array}\right] \quad R_2 \to -R_2$$

Si $k-2=0$ on aura alors trois pivots (car la derniére rangée serait alors $\left[\begin{array}{ccc|c} 0 & 0 & -1 & 1 \end{array}\right]$). Donc aucune rangée contradictoire et aucune colonne sans pivot: une solution unique.

Si $k-2 \neq 0$ on continue:

$$\left[\begin{array}{ccc|c} 1 & 2 & k & 1 \\ 0 & 1 & 1+k & -3 \\ 0 & 0 & -k^2+1 & 3k+3 \end{array}\right] \quad R_3 \to R_3 - (k-2)R_2$$

Si $k=1$ on aurait une rangée contradictoire: aucune solution. Si $k=-1$, on aurait une rangée nulle, et donc une colonne sans pivot: infinité de solutions. Autrement, on a trois pivots, et donc une solution unique.

Résumé: $k=1$ donne aucune solution, $k=-1$ donne une infinité de solutions, tout autre k donne une solution unique.

Terminologie que vous devriez comprendre: opérations de rangée, méthode de Gauss, méthode de Gauss-Jordan, rang.

Mike Newman Université d'Ottawa automne 2008

On veut mieux comprendre la relation entre les systèmes linéaires, et l'engendrance et l'indépendance. On commencera par résoudre un système ayant une colonne de constantes qui sont inconnues. Ensuite, on posera d'autres questions afin de comprendre ce que ce système donne vraiment.

Il n'y a essentiellement rien de nouveau ici. Il s'agit de mettre ensemble les idées qu'on connaît déja, et de voir comment le tout fonctionne ensemble. On verra encore plus de connections quand on parle de matrices.

1. Un Système

Exemple: Donner la forme échelonnée réduite:

$$\left[\begin{array}{ccc|c} 1 & -4 & 7 & r \\ 2 & -5 & 8 & s \\ 3 & -6 & 9 & t \end{array}\right]$$

Voici la réduction:

$$\left[\begin{array}{ccc|c} 1 & -4 & 7 & r \\ 0 & 3 & -6 & -2r+s \\ 0 & 6 & -12 & -3r+t \end{array}\right] \quad R_2 \to R_2 - 2R_1, \quad R_3 \to R_3 - 3R_1$$

$$\left[\begin{array}{ccc|c} 1 & -4 & 7 & r \\ 0 & 3 & -6 & -2r+s \\ 0 & 0 & 0 & r-2s+t \end{array}\right] \quad R_3 \to R_3 - 2R_2$$

$$\left[\begin{array}{ccc|c} 1 & -4 & 7 & r \\ 0 & 1 & -2 & -\frac{2}{3}r+\frac{1}{3}s \\ 0 & 0 & 0 & r-2s+t \end{array}\right] \quad R_2 \to \frac{1}{3}R_2$$

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & -\frac{5}{3}r+\frac{4}{3}s \\ 0 & 1 & -2 & \frac{2}{3}r-\frac{1}{3}s \\ 0 & 0 & 0 & r-2s+t \end{array}\right] \quad R_1 \to R_1 + 4R_2$$

Qu'a-t-on appris? Il y a une rangée nulle dans la matrice des coefficients. Mais dans la matrice augmentée? Ça dépend. Si $0 = r - 2s + t$ on a une solution; sinon, aucune solution. De plus, il y un paramètre, car il manque un pivot dans la troisième colonne. Donc s'il y a une solution, il y aura une infinité de solutions.

Résumé: Si $r - 2s + t \neq 0$, aucune solution. Si $r - 2s + t = 0$, une infinité de solutions.

Mais on a appris beaucoup plus que ça. Voyons quelques exemples de plus.

2. Engendrance

Exemple: Est-ce que les vecteurs $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\}$ engendrent $\mathbb{R}^3$?

En posant

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix} + x_3 \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$$

on obtient directement le système précédent. N'ayant pas de solution en générale, la réponse est non, les vecteurs n'engndrent pas $\mathbb{R}^2$. Il y a des vecteurs (par exemple avec $r - 2s + t \neq 0$!) qui ne peuvent être écrits en combinaison linéaire de ces trois vecteurs. Mais on a appris même plus: la vraie question est *qu'est-ce que* ces trois vecteurs engendrent? Réponse: les vecteurs dans $\mathbb{R}^3$ ayant $r - 2s + t = 0$:

$$\begin{aligned} \operatorname{span} \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\} &= \left\{ \begin{bmatrix} r \\ s \\ t \end{bmatrix} \in \mathbb{R}^3 \mid r - 2s + t = 0 \right\} \\ &= \left\{ s \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} = \operatorname{span} \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} \end{aligned}$$

On vous laisse l'exercice de montrer que ce dernier ensemble est en fait une base. Pour ceux qui veulent un défi, montrer que ce qu'on vient de faire donnera *toujours* une base.

Question: Quelle est la dimension de l'espace engendré? Quel est le nombre de conditions? Quel est le nombre de pivots? Quel est le nombre de paramètres? Quelle est la dimension de $\mathbb{R}^3$? Quels sont les liens entre toutes ces questions?

Exemple: On sait que $\begin{bmatrix} 2 & 1 & 0 \end{bmatrix}^T$ est dans l'espace engendré, car la condition est satisfaite (vérifier svp!). Mais comment l'exprimer en combinaison linéaire de la base qu'on vient de trouver?

On à la solution générale de ce système. Il y a un paramètre, car la troisième colonne ne possède aucun pivot. Mettons $x_3 = p$ (pas t, c'est déja pris!). Donc on a la solution générale:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{5}{3}r + \frac{4}{3}s + p \\ -\frac{2}{3}r + \frac{1}{3}s + 2p \\ p \end{bmatrix} = \begin{bmatrix} -\frac{5}{3}r + \frac{4}{3}s \\ -\frac{2}{3}r + \frac{1}{3}s \\ 0 \end{bmatrix} + p \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Ici, r, s, t sont des constantes. En particulier, $r = 2$, $s = 1$ et $t = 0$. Ce n'est rien d'autre que le vecteur qu'on cherche: $\begin{bmatrix} r & s & t \end{bmatrix}^T = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}^T$. Donc:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{5}{3}(2) + \frac{4}{3}(1) \\ -\frac{2}{3}(2) + \frac{1}{3}(1) \\ 0 \end{bmatrix} + p \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 0 \end{bmatrix} + p \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Pour n'importe quelle valeur de p on aura la solution $x_1 = -2 + p$, $x_2 = -1 + 2p$, $x_3 = p$:

$$(-2 + p) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + (-1 + 2p) \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix} + (p) \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Par exemple et par hasard, prennons $p = 2$; on a $x_1 = 0$, $x_2 = 3$ et $x_3 = 2$:

$$0 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 3 \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix} + 2 \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

3. Indépendance

Exemple: Est-ce que les vecteurs $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\}$ sont indépendants?

En posant une combinaison linéaire égale à $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$ on obtient directement le système précédent avec $r = s = t = 0$. C'est un système homogène. C'est aussi un cas particulier de la solution précédente. La solution générale est

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{5}{3}(0) + \frac{4}{3}(0) + p \\ -\frac{2}{3}(0) + \frac{1}{3}(0) + 2p \\ p \end{bmatrix} = p \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

On a la solution générale $x_1 = p$, $x_2 = 2p$, $x_3 = p$, et donc

$$p \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 2p \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix} + p \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

C'est la solution générale du système homogène. C'est la relation de dépendance la plus générale qui existe entre ces trois vecteurs. Les deux phrases précédentes disent exactement la même chose!

Il y a certainement des solutions non-nulles; on choisit n'importe quel $p \neq 0$. Par exemple et par hasard $p = -3$ donne $x_1 = -3$, $x_2 = -6$ et $x_3 = -3$.

$$-3 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - 6 \begin{bmatrix} -4 \\ -5 \\ -6 \end{bmatrix} - 3 \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Il y a une autre observation à faire: la solution générale du système homogène est exactement une l'espace engendré:

$$\operatorname{span}\left\{\begin{bmatrix}1\\2\\1\end{bmatrix}\right\}$$

Donc la solution générale au système homogène *est* un espace vectoriel, et on connaît une base!

4. Engendrance, Indépendance et Bases

Exemple: Soit U l'espace engendré par ces trois vecteurs (on connaît U maintenant...). Trouver un ensemble engendrant minimal pour U contenu dans l'ensemble original des trois vecteurs.

Attention! on ne cherche pas un ensemble engendrant minimal pour $\mathbb{R}^3$. Ceci est impossible. Pourquoi?

On a trouvé une dépendance, donc on choisit un des vecteurs ayant coefficient non-nul et on l'enlève. Par exemple:

$$\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

On vérifie facilement que c'est une base pour U, car c'est indépendant.

Exemple: Trouver un ensemble indépendant maximal dans $\mathbb{R}^3$ contenant les deux vecteurs de votre ensemble engendrant minimal pour U.

Note que les deux vecteurs engendrent U aussi (pourquoi?), donc on cherche un vecteur qui n'est pas dans U. On le fait directement de la condition. On choisit s et t n'importe comment et on choisit $r \neq 2s - t$. On ajoute ce vecteur.

Par exemple et par hasard, prennons $s = 1$, $t = 2$. Donc $2s - t = 0$ et il faut choisir $r \neq 0$. Par exemple et par hasard, prennons $r = 1$. Donc $\begin{bmatrix}1 & 1 & 2\end{bmatrix}^T$ n'est pas dans U, et on a l'ensemble indépendant:

$$\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}7\\8\\9\end{bmatrix}, \begin{bmatrix}1\\1\\2\end{bmatrix}\right\}$$

On vérifie facilement que c'est une base pour $\mathbb{R}^3$, car on connaît la dimension de $\mathbb{R}^3$.

Question: On a dû faire des vrais calculs afin de trouver la forme échelonnée réduite. Mais après cette étape, qu'est-ce qu'on a dû *calculer*? Qu'est-ce qu'on a pu *lire*? Est-ce qu'on avait vraiment besoin de calculer la forme échelonnée réduite, ou est-ce que la forme échelonnée aurait suffit?

5. Un Exercice

Exercice: Considérer la petite réduction suivante:

$$\left[\begin{array}{cccccc|c} 0 & 1 & 1 & -1 & 1 & 2 & a \\ 1 & 0 & 1 & 2 & 2 & -2 & b \\ 0 & -2 & -2 & 2 & 1 & -7 & c \\ 2 & 3 & 5 & 1 & 7 & 2 & d \\ 3 & 0 & 3 & 6 & 3 & -3 & e \end{array}\right] \qquad \to \cdots \to \qquad \left[\begin{array}{cccccc|c} 1 & 0 & 1 & 2 & 2 & -2 & b \\ 0 & 1 & 1 & -1 & 1 & 2 & a \\ 0 & 0 & 0 & 0 & 1 & -1 & (2a+c)/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3a+2b+d \\ 0 & 0 & 0 & 0 & 0 & 0 & 2a-3b+e \end{array}\right]$$

À vous: donner l'espace engendré. Décrire la dépendance des colonnes. Donner une base pour la solution générale du système homogène (attention: il faudrait montrer que votre "base" est indépendante). Trouver un ensemble engendrant minimal pour l'espace engendré. Trouver un ensemble indépendent maximal pour $\mathbb{R}^3$ (du moins trouver *un* vecteur que vous pouvez ajouter pour former un ensemble indépendant plus grand).

Il n'y a (presque!) aucun calcul à faire pour répondre à ces questions. Il s'agit d'interpréter la forme échelonnée.

Au cas où vous voudriez savoir la forme échelonnée réduite, demandez-vous si il suffirait de la connaître pour le système *homogène* ($a = b = c = d = e = 0$). Ce serait alors un peu plus facile...

On commence avec une définition et quelques exemples. On verra le "pourquoi" dans la prochaine section!

1. Définition

Pour une matrice A, on écrit souvent A_{ij} pour l'élément à la position (i, j), c'est-à-dire dans la i-ième rangée et la j-ième colonne de A. Donc si $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ alors $A_{2,2} = 5$ et $A_{1,2} = 2$; $A_{3,1}$ et $A_{1,4}$ n'existent pas.

Définition 1. *Soit A une matrice de taille $m \times k$ et B une matrice de taille $k \times n$.*

Le produit AB est une matrice de taille $m \times n$, avec $(AB)_{ij}$ égal au produit scalaire de la i-ième rangée de A avec la j-ième colonne de B (considérés comme vecteurs-colonnes).

Autrement dit, $(AB)_{ij} = \sum_{t=1}^{k} (A)_{it} B_{tj}$.

Voyons un example.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 3 & 0 & 3 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad AB = \begin{bmatrix} 1 & 2 & 3 \\ \boxed{1} & \boxed{0} & \boxed{\text{-1}} \end{bmatrix} \begin{bmatrix} \boxed{3} & 0 & 3 \\ \boxed{1} & 1 & 1 \\ \boxed{0} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 8 \\ \boxed{3} & 0 & 2 \end{bmatrix}$$

Comme exemple, on voit le calcul de la position $(2, 1)$ du produit. Il s'agit de prendre la première rangée de A, et la première colonne de B, et de calculer leur produit scalaire afin de déterminer la valeur dans la position $(2, 3)$ du produit:

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = (1)(3) + (0)(1) + (-1)(0) = 3 + 0 + 0 = 3$$

Exercice: Vérifier les autres éléments de ce produit, selon ce même règle.

Pour une multiplication matricielle, le nombre de *colonnes* dans la première matrice doit être égal au nombre de *rangées* dans la deuxième. Le résultat de AB possède le même nombre de rangées que A et le même nombre de colonnes que B.

Exemple: Lesquels des suivants sont compatibles? Quelle est la taille du produit?

$$\begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}4\\5\\6\end{bmatrix} \qquad \begin{bmatrix}4\\5\\6\end{bmatrix}\begin{bmatrix}1 & 2 & 3\end{bmatrix} \qquad \begin{bmatrix}r & s & t\\1 & 2 & 3\end{bmatrix}\begin{bmatrix}4\\5\end{bmatrix}$$

$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}\begin{bmatrix}1 & 2 & 3 & 4 & 5\\0 & 1 & 0 & 1 & 0\\2 & 4 & 6 & 8 & 0\end{bmatrix} \qquad \begin{bmatrix}a & 4\\b & 5\\c & 6\end{bmatrix}\begin{bmatrix}4\\5\end{bmatrix}+\begin{bmatrix}8\\9\end{bmatrix}$$

Le premier: $1 \times \mathbf{3}$ est compatible avec $\mathbf{3} \times 1$, car $\mathbf{3} = \mathbf{3}$. Le résultat serait 1×1.

Ensuite, $3 \times \mathbf{1}$ est compatible avec $\mathbf{1} \times 3$, car $\mathbf{1} = \mathbf{1}$. Le résultat serait 3×3.

Ensuite, $2 \times \mathbf{3}$ n'est pas compatible avec $\mathbf{2} \times 1$, car $\mathbf{3} \neq \mathbf{2}$.

Ensuite, $2 \times \mathbf{3}$ est compatible avec $\mathbf{3} \times 5$, car $\mathbf{3} \neq \mathbf{3}$. Le résultat serait 2×5.

Ensuite $3 \times \mathbf{2}$ est compatible avec $\mathbf{2} \times 1$, car $\mathbf{2} = \mathbf{2}$, pour donner 3×1. Mais on ne peut pas additionner 3×1 et 2×1, donc la somme n'est pas compatible.

Note que les *tailles* qui déterminent si un produit matriciel existe; les valeurs ne jouent aucun rôle. Elles jouent un rôle pour calculer le produit!

Exercice: Calculer les produits des matrices dans l'exemple précédent. Ceux qui existent, bien sûr!

NB: Une matrice de taille $n \times 1$ est la même chose qu'un vecteur-colonne dans $\mathbb{R}^n$. Une matrice de taille 1×1 est (essentiellement) la même chose qu'un chiffre.

2. Combinaisons Linéaires, Systèmes d'Équations

Supposons que $c_1, c_2, \ldots, c_n$ sont des vecteurs dans $\mathbb{R}^m$; c'est-à-dire que chaque $\mathbf{c}_i$ est une colonne de m chiffres. Supposons de plus que $a_1, a_2, \ldots, a_n$ sont des chiffres réels. On a (souvent!) vu une combinaison linéaire:

$$a_1\mathbf{c}_1 + a_2\mathbf{c}_2 + \cdots + a_n\mathbf{c}_n$$

Formons une matrice A del que les colonnes de A sont $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$. On écrit souvent que $A = \left[\mathbf{c}_1|\mathbf{c}_2|\ldots|\mathbf{c}_n\right]$. Aussi, mettons u pour le vecteur-colonne ayant les éléments $a_1, a_2, \ldots, a_n$. Alors

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_=\mathbf{u}_= A\mathbf{u}$$

La multiplication matricielle est une forme compacte des combinaison linéaires!

Voyons un exemple. Soit

$$\mathbf{c}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix} \qquad \mathbf{c}_1 = \begin{bmatrix}2\\-1\\3\end{bmatrix} \qquad \mathbf{c}_1 = \begin{bmatrix}0\\1\\2\end{bmatrix} \qquad \mathbf{c}_1 = \begin{bmatrix}2\\-1\\5\end{bmatrix}$$

et $a_1 = 1$, $a_2 = 4$, $a_3 = -2$, $a_4 = -3$. Alors

$$\begin{aligned} a_1\mathbf{c}_1 + a_2\mathbf{c}_2 + a_3\mathbf{c}_3 + a_4\mathbf{c}_4 &= (1)\begin{bmatrix}1\\0\\1\end{bmatrix} + (4)\begin{bmatrix}2\\-1\\3\end{bmatrix} + (-2)\begin{bmatrix}0\\1\\2\end{bmatrix} + (-3)\begin{bmatrix}2\\-1\\5\end{bmatrix} \\ &= \begin{bmatrix}1\\0\\1\end{bmatrix} + \begin{bmatrix}8\\-4\\12\end{bmatrix} + \begin{bmatrix}0\\-2\\-4\end{bmatrix} + \begin{bmatrix}-6\\3\\-15\end{bmatrix} \\ &= \begin{bmatrix}3\\-3\\-6\end{bmatrix} \end{aligned}$$

Par contre, on aurait alors

$$A\mathbf{u} = \begin{bmatrix}1 & 2 & 0 & 2\\0 & -1 & 1 & -1\\1 & 3 & 2 & 5\end{bmatrix}\begin{bmatrix}1\\4\\-2\\-3\end{bmatrix} = \begin{bmatrix}3\\-3\\-6\end{bmatrix}$$

C'est très générale!

C'est une interprétation de la multiplication en termes des *colonnes* de la matrice A. On peut aussi penser en termes des *rangées* de A.

Soit A la matrice des coefficients d'un système d'équations linéaires, et $\mathbf{b}$ la colonne des constantes. Donc la matrice augmentée est $[A|b]$. Soit $\mathbf{x}$ le vecteur des variables; $\mathbf{x}$ est un vecteur-colonne dans $\mathbb{R}^n$. Alors le système d'équations originale est exactement $A\mathbf{x} = \mathbf{b}$!

Voyons un exemple. Voici un système d'équations:

$$x_1 + 3x_2 + x_3 = 5 \qquad 2x_1 + x_3 + 2x_4 = 0 \qquad 5x_1 + 4x_2 + x_4 = 2$$

et voici sa matrice augmentée:

$$\left[\begin{array}{cccc|c}1 & 3 & 1 & 0 & 5\\2 & 0 & 1 & 2 & 0\\5 & 4 & 0 & 1 & 2\end{array}\right]$$

Chaque rangée correspond à une équation, et les colonnes correspondent aux variables. L'équation matriciel $A\mathbf{x} = \mathbf{b}$ donne

$$\begin{bmatrix}1 & 3 & 1 & 0\\2 & 0 & 1 & 2\\5 & 4 & 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}5\\0\\2\end{bmatrix}$$

$$\begin{bmatrix}x_1 + 3x_2 + x_3\\2x_1 + x_3 + 2x_4\\5x_1 + 4x_2 + x_4\end{bmatrix} = \begin{bmatrix}5\\0\\2\end{bmatrix}$$

C'est une égalité entre deux vecteurs-colonnes, voulant dire que chaque élément correpondant doit être égal:

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 5 \\ 2x_1 + x_3 + 2x_4 &= 0 \\ 5x_1 + 4x_2 + x_4 &= 2 \end{aligned}$$

Donc "$A\mathbf{x} = \mathbf{b}$" équivaut au système! La multiplication matriciel est une forme compacte d'un système d'équations.

3. Propriétés

On veut étudier l'arithméthique des matrices. On connaît déja l'addition et la multiplication scalaire; maintenant on y ajoute la multiplication matricielle. Les suivants sont valides pour toutes matrices A, B et C telles que les expressions existent.

Exercice: En lisant la liste des propriétés, déterminer les contraintes sur les tailles des matrices pour que l'expression existe.

(1) $\boxed{A(B+C) = AB + AC}$

(2) $\boxed{(A+B)+C = AC + BC}$

(3) $\boxed{(AB)C = A(BC)}$

Voici une preuve pour les matrices de taille 2×2:

$$\begin{aligned} \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix}\right) \begin{bmatrix} i & j \\ k & l \end{bmatrix} &= \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix} \begin{bmatrix} i & j \\ k & l \end{bmatrix} \\ &= \begin{bmatrix} (ae+bg)i + (af+bh)k & (ae+bg)j + (af+bh)l \\ (ce+dg)i + (cf+dh)k & (ce+dg)j + (cf+dh)l \end{bmatrix} \\ &= \begin{bmatrix} aei + bgi + +afk + bhk & aej + bgj + afl + bhl \\ cei + dgi + cfk + dhk & cej + dgj + cfl + dhl \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\begin{bmatrix} e & f \\ g & h \end{bmatrix} \begin{bmatrix} i & j \\ k & l \end{bmatrix}\right) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} ei+fk & ej+fl \\ gi+hk & gj+hl \end{bmatrix} \\ &= \begin{bmatrix} a(ei+fk) + b(gi+hk) & a(ej+fl) + b(gj+hl) \\ c(ei+fk) + d(gi+hk) & c(ej+fl) + d(gj+hl) \end{bmatrix} \\ &= \begin{bmatrix} aei + afk + bgi + bhk & aej + afl + bgj + bhl \\ cei + cfk + dgi + dhk & cej + cfl + dgj + dhl \end{bmatrix} \end{aligned}$$

Les deux expressions étant égal, l'énoncé est prouvé pour les matrices des taille 2×2.

Voici une preuve pour les matrices en générale. C'est plus courte, grâce à la notation Σ. On calcule la valeur à la position (i, j) pour chaque côté:

$$((AB)C)_{ij} = \sum_{t=1}^{n} (AB)_{it} C_{tj} = \sum_{t=1}^{n} \left(\sum_{s=1}^{m} A_{is} B_{st} \right) C_{tj} = \sum_{t=1}^{n} \sum_{s=1}^{m} A_{is} B_{st} C_{tj}$$

$$(A(BC))_{ij} = \sum_{p=1}^{m} A_{ip} (BC)_{pj} = \sum_{p=1}^{m} A_{ip} \left(\sum_{q=1}^{n} B_{pq} C_{qj} \right) = \sum_{p=1}^{m} \sum_{q=1}^{n} A_{ip} B_{pq} C_{qj}$$

Les deux expressions sont égales; t et s correspondent q et p. Donc l'énoncé est prouvé en générale. Question: qu'est-ce que c'est m et n? Indice: si les produits existent, il y a des contraintes sur les tailles de A, B et C.

(4) $\boxed{k(AB) = (kA)B = A(kB) = (AB)k, \text{ pour } k \in \mathbb{R}}$

On définit la matrice $\mathbf{0}_{mn}$ comme la matrice de taille $m \times n$ ayant tout élément zéro. Par exemple:

$$\mathbf{0}_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \mathbf{0}_{3,1} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

C'est le zéro de l'espace vectoriel. D'habitude on écrit seulement $\mathbf{0}$ et la taille est inférée du contexte.

(5) $\boxed{A\mathbf{0} = \mathbf{0} \text{ et } \mathbf{0}B = \mathbf{0}}$

On définit la matrice I_n comme la matrice de taille $n \times n$ ayant 1 sur la diagonale et 0 ailleurs. Note que c'est toujours carrée (nombre de rangées égal nombre de colonnes). Par exemple:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad I_1 = [1] = 1$$

D'habitude on écrit seulement I et la taille est inférée du contexte. Cette matrice est une identité pour la multiplication (on dit "I" pour identité):

(6) $\boxed{AI = A \text{ et } IB = B}$

Pour toute matrice A, on définit la TRANSPOSE de A comme étant la matrice avec les rangées et colonnes interchangées. Par exemple:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \quad \begin{bmatrix} a & b & c \end{bmatrix}^T = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

On revoit une notation qu'on avait vu au début du cours afin d'écrire des vecteurs de façon plus compact sur la page.

(7) $\boxed{(A^T)^T}$

(8) $\boxed{(AB)^T = B^T A^T}$

Cette dernière est un peu différent: en prennant la transpose d'un produit, on renverse l'ordre? Vérifions les tailles. Si A est $m \times k$ et B est $k \times n$, alors AB est $m \times n$ et $(AB)^T$ est $n \times m$. À droite, B^T est de taille $n \times k$ et A^T est de taille $k \times m$, donc $B^T A^T$ est de taille $n \times m$.

Exercice: Soit $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ et $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$. Calculer $(AB)^T$ et $B^T A^T$, et alors montrer directement que $(AB)^T = B^T A^T$.

Il semble que l'arithmétique des matrices est essentiellement la même chose que l'arithmétique de $\mathbb{R}$. Ce n'est pas vrai, il reste des surpries:

(1) $\boxed{\text{Parfois } AB \text{ existe mais pas } BA.}$

On a déja vu ceci: par exemple A 2×3 et B 3×5.

(2) $\boxed{\text{Parfois } AB \text{ et } BA \text{ existent mais } AB \neq BA.}$

Exercice: Montrer que $AB \neq BA$ pour $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ et $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

(3) $\boxed{\text{Parfois } AB = \mathbf{0} \text{ mais } A \neq \mathbf{0} \text{ et } B \neq \mathbf{0}.}$

On voit que $\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. De plus, la multiplication dans l'autre ordre existe, mais ne donne pas zéro! En réelles, si $ax = 0$ et $a \neq 0$, on peut conclure que $x = 0$. Ce n'est pas vrai pour des matrices!

(4) $\boxed{\text{Parfois } A^2 = \mathbf{0} \text{ mais } A \neq \mathbf{0}.}$

On voit que $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Exercice: Trouver une matrice A tel que $A \neq \mathbf{0}$, $A^2 \neq \mathbf{0}$, $A^3 = \mathbf{0}$. Pour tout k, trouver une matrice tel que $A^j \neq \mathbf{0}$ pour $j = 1, 2, \ldots, k-1$ mais $A^k = \mathbf{0}$.

(5) $\boxed{\text{Parfois } AC = BC \text{ mais } A \neq B.}$

On a l'exemple suivant:

$$\begin{bmatrix} 0 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{mais} \qquad \begin{bmatrix} 0 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

Calculer les deux côtés pour prouver l'égalité! Ceci veut dire qu'on ne ne peut pas "canceler" une matrice. Dans les réelles, si on a $rs = ts$ pour $s \neq 0$, on peut diviser par s pour conclure que $r = t$. Ce n'est pas vrai pour des matrices! La raison fondamentale pour cette difficulté (et la surprise (4) aussi) est qu'on n'a pas une division pour les matrices. Il existe (parfois!) des inverses (on les verra dans quelques jours), mais il existe beaucoup de matrices non-zéro qui n'ont pas d'inverses.

Exemple: Simplifier $(A + B)(A - B)$. On voit que

$$(A + B)(A - B) = A^2 - AB + BA - B^2.$$

On ne peut rien faire de plus! Noter qu'il faut faire attention à l'ordre du produit.

Exercice: Donner un exemple de matrices A et B tel que $(A+B)^2 \neq A^2+2AB+B^2$. Indice: on a déja vu un exemple pas loin!

Exemple: Sachant que $C \neq \mathbf{0}$, résoudre $(A - B)(C + D) = -BD + AD$. On imaginera de faire

$$\begin{aligned} (A - B)(C + D) &= -BD + AD \\ AC + AD - BC - BD &= -BD + AD \\ AC + AD - BC - BD + BD - AD &= \mathbf{0} \\ AC - BC &= \mathbf{0} \end{aligned}$$

On ne peut plus rien simplifier. En particulier, on ne peut *pas* diviser par C pour conclure que $A = C$.

MAT 1741 15a. Interprétations Géométriques de Solutions et d'Espaces

Mike Newman | Université d'Ottawa | automne 2008

Ce chapitre est partiellement optionnelle. On verra des exemples plus concrets plus tard. Les deux théoèmes sont importants!

1. Équations de Droites et Plans

On se souvient de la forme de l'équation d'une droite: on a un point et en vecteur directeur. Pour un plan, on a un point et deux vecteurs directeurs. La solution générale pour un système d'équations généralise: le "point" c'est une solution particulière, les "vecteurs directeurs" c'est une base pour le noyau. On avait vu qu'il y a beauvoup d'équations qui correspondent au même droite: prendre un point différent sur la droite, ou un multiple du vecteur directeur. En termes de solutions, on voit que c'est pareil: n'importe quelle solution pariculière peut servir comme $\mathbf{x}^{(P)}$, et il existe typiquement un nombre infini de bases pour le noyau. La méthode de Gauss-Jordan permet d'en trouver un cas particulier de manière efficace.

Est-ce qu'un plan est un sous-espace de $\mathbb{R}^m$? On a vu ce genre de question déja: la réponse est oui, si le plan passe par l'origine. Il en est de même pour la solution générale: c'est un sous-espace exactement lorsqu'elle inclut l'origine. Mais si elle inclut l'origine, alors $\mathbf{x} = 0$ est une solution, voulant dire que $\mathbf{b} = \mathbf{0}$.

Théorème 1. *La solution générale pour* $A\mathbf{x} = \mathbf{b}$ *est un sous-espace si est seulement si le système est homogène, en autre mots que* $\mathbf{b} = \mathbf{0}$

On a vu ce résultat dans plusieurs exemples, exercices, devoirs...

Exercice: Est-ce que chacun est sous-espace? Écrire un système d'équations qui décrit l'ensemble, et donner sa solution générale.

- $\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + z = 1 \right\}$ dans $\mathbb{R}^3$?
- $\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + z = 0 \right\}$ dans $\mathbb{R}^3$?
- $\{p(x) \in \mathbb{P}_2 \mid p(3) = 1\}$ dans $\mathbb{P}_3$?
- $\{p(x) \in \mathbb{P}_2 \mid p(3) = 0\}$ dans $\mathbb{P}_3$?

- $\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a+d=1 \right\}$ dans $\mathbb{M}_{2,2}$?
- $\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a+d=0 \right\}$ dans $\mathbb{M}_{2,2}$?

2. Directions Permises et Directions Défendues

On a une autre interprétation: on se souvient que les équations linéaires forment un espace vectoriel on peut les additionner, etc. On a aussi que les équations d'un système homogène sont exactement les rangées de la matrice des coefficients. Donc l'espace des équations d'un système homogène est exactement à $\mathrm{col}(A^T)$

De l'autre côté, tout vecteur dans $\ker(A)$ correspond à une solution au système homogène. Donc, c'est une solution à tout combinaison des équations dans ce système. Autrement dit, tout vecteur dans $\ker(A)$ est *orthogonal* à tout vecteur dans $\mathrm{col}(A^T)$. Les deux espaces sont orthogonaux!

Théorème 2. $\mathrm{col}(A^T)$ *est orthogonal à* $\ker(A)$.

On a encore une interprétation géométrique: un plan dans $\mathbb{R}^3$ peut se décrire avec *un* vecteur normal ou *deux* vecteurs directeurs. Ce plan correspond à l'espace rangée d'une matrice A: la base pour $\mathrm{col}(A^T)$ correspond aux deux vecteurs directeurs, la base pour $\ker(A)$ correspond au vecteur normal.

Quelle est la taille de cette matrice? Elle aura trois colonnes, puisque l'espace rangée (eg, le plan!) est sous-espace de $\mathbb{R}^3$. Il y aura une colonne sans pivot pour donner $\dim(\ker(A)) = 1$. Il y aura donc deux pivots (dans la forme échelonnée).

Exercice: Déterminer une base pour $\ker(A)$ pour $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$. Monter alors que c'est un plan dans $\mathbb{R}^3$. Tenter de trouver une base pour $\mathrm{col}(A^T)$.

Si c'est vraiment un plan, on aura $\dim(\ker(A)) = 1$ (un vecteur normal) et aussi $\dim(\mathrm{col}(A^T)) = 2$ (deux vecteurs directeurs), donnant $1+2=3$, dans le sens de $\mathbb{R}^3$. Il semble que:

$$\dim(\ker(A)) + \mathrm{rang}(A^T) \;\; ?=? \;\; n$$

C'est très similaire à un théorème qu'on a vu! Mais est-ce vrai? On verra bientôt, mais pour le moment, c'est un excellent exercice!

1. Quatre Espaces

On identifie quatre espaces pour une matrice A.

Définition 1. *Soit A une matrice de taille $m \times n$.*

- $\mathrm{col}(A)$ *est l'espace engendré par les colonnes de A.*
- $\ker(A)$ *est l'espace de toutes les solutions à $A\mathbf{x} = \mathbf{0}$.*
- $\mathrm{col}(A^T)$ *est l'espace engendré par les rangées de A.*
- $\ker(A^T)$ *est l'espace de toutes les solutions à $A^T\mathbf{y} = \mathbf{0}$.*

On dit ESPACE COLONNE DE A pour $\mathrm{col}(A)$, ESPACE RANGÉE DE A pour $\mathrm{col}(A^T)$ et NOYAU DE A pour $\ker(A)$ (l'espace $\ker(A^T)$ n'a pas de nom spécial).

En symboles on a:

Définition 2. *Soit A une matrice de taille $m \times n$.*
Soit $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ les colonnes de A, et $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ les rangées de A.

- $\mathrm{col}(A) = \mathrm{span}\,\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\} \subseteq \mathbb{R}^m$
- $\ker(A) = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^n$
- $\mathrm{col}(A^T) = \mathrm{span}\,\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\} \subseteq \mathbb{R}^n$
- $\ker(A^T) = \left\{\mathbf{y} \mid A^T\mathbf{y} = \mathbf{0}\right\} \subseteq \mathbb{R}^m$

On voit que l'espace colonne $\mathrm{col}(A)$ est un espace engendré par un ensemble de vecteurs. Donc c'est réellement un espace vectoriel: c'est le théorème qui dit que $\mathrm{span}\,\{\cdots\}$ est toujours un espace vectoriel. De plus un sous-espace ce $\mathbb{R}^m$. De même, on voit que l'espace rangée $\mathrm{col}(A^T)$ est un sous espace de $\mathbb{R}^n$.

On vérifie, à l'aide du test de sous-espace, que $\ker(A)$ est un sous-espace de $\mathbb{R}^n$.

(1) $A\mathbf{0} = \mathbf{0}$; c'est un des propriétés de la multiplication matricielle, donc $\mathbf{0}$ est dans $\ker(A)$.
(2) Si $\mathbf{u}$ et $\mathbf{v}$ sont dans $\ker(A)$, alors $A\mathbf{u} = A\mathbf{v} = \mathbf{0}$. Donc $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u}+A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$ par des propriétés de multiplication matricielle. Alors $\mathbf{u} + \mathbf{v}$ est dans $\ker(A)$.

(3) Si $\mathbf{u}$ est dans $\ker(A)$ et $k \in \mathbb{R}$ alors $A\mathbf{u} = \mathbf{0}$. Donc $A(k\mathbf{u}) = k(A\mathbf{u}) = k(\mathbf{0}) = \mathbf{0}$; encore des propriétés de la multiplication des matrices. Alors $k\mathbf{u}$ est dans $\ker(A)$.

La preuve est similaire pour $\ker(A^T)$, qui est sous-espace de $\mathbb{R}^m$.

L'espace $\ker(A)$ est en fait une chose qu'on a souvent vu, car $A\mathbf{x} = \mathbf{0}$ équivaut à un ensemble de conditions. Le produit $A\mathbf{x}$ est une matrice de taille $m \times 1$. Chaque élément de cette matrice est égal à zéro, donc on retrouve une équation linéaire avec constante 0, pour chaque rangée de A.

Exemple: Soit $A = \begin{bmatrix} 1 & 0 & 3 & -2 \\ -1 & 2 & 1 & 1 \\ -1 & 4 & 5 & 0 \end{bmatrix}$. Alors

- $\mathrm{col}(A) = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$

- $$\ker(A) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mid \begin{bmatrix} 1 & 0 & 3 & -2 \\ -1 & 2 & 1 & 1 \\ -1 & 4 & 5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$
  $$= \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \mid \begin{array}{r} x_1 + 3x_3 - 2x_4 = 0 \\ -x_1 + 2x_2 + x_3 + x_4 = 0 \\ -x_1 + 4x_2 + 5x_3 = 0 \end{array} \right\}$$

- $\mathrm{col}(A^T) = \mathrm{span}\left\{ \begin{bmatrix} 1 & 0 & 3 & -2 \end{bmatrix}, \begin{bmatrix} -1 & 2 & 1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 4 & 5 & 0 \end{bmatrix} \right\}$

- $$\ker(A^T) = \mathrm{span}\left\{ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \mid \begin{bmatrix} 1 & -1 & -1 \\ 0 & 2 & 4 \\ 3 & 1 & 5 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$
  $$= \left\{ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \mid \begin{array}{r} y_1 - y_2 - y_3 = 0 \\ 2y_2 + 4y_3 = 0 \\ 3y_1 + y_2 + 5y_3 = 0 \\ -2y_1 + y_2 = 0 \end{array} \right\}$$

2. Espaces et Solutions

On observe que le noyau est exactement la solution générale au système homogène avec matrice de coefficients A. Donc on connaît une méthode pour déterminer $\ker(A)$: la méthode de Gauss-Jordan.

Exemple: Trouver $\ker(A)$ pour l'exemple ci-haut.

$$\begin{bmatrix} 1 & 0 & 3 & -2 \\ -1 & 2 & 1 & 1 \\ -1 & 4 & 5 & 0 \end{bmatrix} \xrightarrow[R_3 \to R_3 + R_1]{R_2 \to R_2 + R_1} \begin{bmatrix} 1 & 0 & 3 & -2 \\ 0 & 2 & 4 & -1 \\ 0 & 4 & 8 & -2 \end{bmatrix}$$
$$\xrightarrow[R_2 \to \frac{1}{2} R_2]{} \begin{bmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & 2 & -\frac{1}{2} \\ 0 & 4 & 8 & -2 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 4R_2]{} \begin{bmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solution générale est (x_3 et x_4 sont les deux paramètres)

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3x_3 + 2x_4 \\ -2x_3 + x_4/2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 1/2 \\ 0 \\ 1 \end{bmatrix}$$

Alors le noyau est

$$\ker(A) = \left\{ x_3 \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 1/2 \\ 0 \\ 1 \end{bmatrix} \mid x_3, x_4 \in \mathbb{R} \right\}$$

C'est un sous-espace de $\mathbb{R}^4$ engendré par deux vecteurs. Est-ce que ces deux vecteurs sont inépendents? Essayons de faire une combinaison linéaire égale à zéro:

$$p \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \end{bmatrix} + q \begin{bmatrix} 2 \\ 1/2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Ceci donne quatre équations:

$$-3p + 2q = 0 \qquad -2p + q/2 = 0 \qquad p = 0 \qquad q = 0$$

Chaque équation correspond à une des variables x_1, x_2, x_3, x_4. Si on regarde aux équations qui correspondent aux paramètres x_3 et x_4, on voit $p = q = 0$. Donc les deux vecteurs sont indépendants. On a souvent vu ceci dans les exercices, mais c'est le momwent d'observer que c'est un résultat générale:

Théorème 3. *Pour toute matrice A, la solution générale de $A\mathbf{x} = \mathbf{0}$, qu'on lit de la forme échelonnée réduite, donne toujours une base pour* $\ker(A)$.

Ce n'est pas nécessaire de prendre la forme échelonnée réduite. L'essentiel est de connaître tout les pivots, donc une forme échelonnée suffit. Mais la forme échelonnée réduite est plus simple pour lire la solution.

La dimension de $\ker(A)$ est le nombre de colonnes *sans* pivot. Le rang de A est le nomber de colonnes *avec* pivot. On a donc un lien fondamentale entre les espaces:

Théorème 4. *Pour toute matrice A de taille $m \times n$, on a* $\operatorname{rang}(A) + \dim(\ker(A)) = n$

Exercice: Montrer que si A est de taille $m \times n$, on a $\operatorname{rang}(A^T) + \dim(\ker(A^T)) = m$.

Le système homogène ce montre donc important. Mais il y a un autre lien. Soit $A\mathbf{x} = \mathbf{b}$ un système linéaire, avec $\mathbf{x}^{(P)}$ une solution paritucière, et $\mathbf{x}^{(H)}$ une solution au système homogène $A\mathbf{x} = \mathbf{0}$. Alors

$$A\left(\mathbf{x}^{(P)} + \mathbf{x}^{(H)}\right) = A\mathbf{x}^{(P)} + A\mathbf{x}^{(H)} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Donc $\mathbf{x}^{(P)} + \mathbf{x}^{(H)}$ est aussi une solution pour $A\mathbf{x} = \mathbf{b}$.

Exemple: Donner la solution générale pour la matrice augmentée $\left[\begin{array}{cccc|c} 1 & 0 & 3 & -2 & 1 \\ -1 & 2 & 1 & 1 & 3 \\ -1 & 4 & 5 & 0 & 7 \end{array}\right]$

$$\left[\begin{array}{cccc|c} 1 & 0 & 3 & -2 & 1 \\ -1 & 2 & 1 & 1 & 3 \\ -1 & 4 & 5 & 0 & 7 \end{array}\right] \xrightarrow[R_3 \to R_3 + R_1]{R_2 \to R_2 + R_1} \left[\begin{array}{cccc|c} 1 & 0 & 3 & -2 & 1 \\ 0 & 2 & 4 & -1 & 4 \\ 0 & 4 & 8 & -2 & 8 \end{array}\right]$$

$$\xrightarrow[R_2 \to \frac{1}{2} R_2]{} \left[\begin{array}{cccc|c} 1 & 0 & 3 & -2 & 1 \\ 0 & 1 & 2 & -\frac{1}{2} & 2 \\ 0 & 4 & 8 & -2 & 8 \end{array}\right] \xrightarrow[R_3 \to R_3 - 4R_2]{} \left[\begin{array}{cccc|c} 1 & 0 & 3 & -2 & 1 \\ 0 & 1 & 2 & -\frac{1}{2} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

On a la solution générale

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 1/2 \\ 0 \\ 1 \end{bmatrix}$$

C'est presque la même chose que l'exemple précédent. Les opérations de rangée sont identiques. La matrice des coefficients est pareil en forme échelonnée réduite. Les seules différences sont à droite de la barre. Donc la solution est pareil, *sauf pour le vecteur constant*. La solution générale est exactement la somme d'une solution particulière et la solution générale homogène.

$$\mathbf{x}^{(P)} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{x}^{(H)} = x_3 \begin{bmatrix} -3 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 1/2 \\ 0 \\ 1 \end{bmatrix}$$

Théorème 5. *Soit* $A\mathbf{x} = \mathbf{b}$ *un système d'équations linéaires.*

Soit $\mathbf{x}^{(P)}$ *une solution particulière et* $\mathbf{x}^{(H)}$ *la solution générale au système homogène* $A\mathbf{x} = \mathbf{0}$.

Alors la solution générale pour $A\mathbf{x} = \mathbf{0}$ *est* $\mathbf{x}^{(P)} + \mathbf{x}^{(H)}$

Supposons que $A\mathbf{x} = \mathbf{b}$ est consistant. Alors $A\mathbf{x} = \mathbf{b}$ possède une solution *unique* si et seuelement si $A\mathbf{x} = \mathbf{0}$ possède une solution unique. Si $\mathbf{u}$ et $\mathbf{v}$ sont deux solutions particulières distinctes, alors leur différence $\mathbf{u} - \mathbf{v}$ est une solution non-triviale au système homogène, car $A(\mathbf{u} - \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{b} - \mathbf{b} = \mathbf{0}$.

3. Quatre Théorèmes

On connaît beaucoup de résultats entre les solutions, les matrices, les équations... C'est le moment de mettre le tout ensemble. Il n'y a essentiellement rien de nouveau dans ces quatre théorèmes. Il s'agit de combiner les résultats qu'on connaît déja.

Dans chaque théorème, A est une matrice de taille $m \times n$, $\mathbf{x}$ est un vecteur-colonne dans $\mathbb{R}^n$ et $\mathbf{b}$ est un vecteur-colonne dans $\mathbb{R}^m$. Comme d'habitude!

Théorème 6.

$A\mathbf{x} = \mathbf{b}$ *est consistant*

$\iff \mathbf{b} \in \operatorname{col}(A)$

$\iff \mathbf{b}$ *est engendré par les colonnes de* A

$\iff \mathbf{b}$ *est combinaison linéaire des colonnes de* A

$\iff$ *aucune rangée contradictoire dans la forme échelonnée de* $[A|b]$

$\iff \operatorname{rang}(A|b) = \operatorname{rang}(A)$

On peut alors déterminer si un vecteur est engendré par un ensemble donné: l'ensemble devient les colonnes d'une matrice, le vecteur devient la colonne des constantes, et on solutionne.

Théorème 7.

$$\begin{aligned}
&A\mathbf{x} = \mathbf{0} \textit{ possède une solution unique} \\
&\quad \iff A\mathbf{x} = \mathbf{0} \textit{ implique } \mathbf{x} = \mathbf{0} \\
&\quad \iff \mathbf{0} \textit{ n'est pas combinaison linéaire non-triviale des colonnes de } A \\
&\quad \iff \textit{les colonnes de } A \textit{ sont linéairement indépendants} \\
&\quad \iff \textit{aucun paramètre dans la forme échelonnée de } A \\
&\quad \iff \operatorname{rang}(A) = n
\end{aligned}$$

On peut alors déterminer un ensemble indépendant des colonnes de A: on retient les colonnes qui correspondent aux pivots! Ce serait un ensemble indépendant maximal dans $\operatorname{col}(A)$. Pourquoi? Car si on ajoute un autre vecteur, on aurait une colonne sans pivot. Donc on connaît comment trouver une base pour $\operatorname{col}(A)$. C'est à revoir

Théorème 8.

$$\begin{aligned}
&A\mathbf{x} = \mathbf{b} \textit{ est consistant pour tout } \mathbf{b} \in \mathbb{R}^m \\
&\quad \iff \mathbf{b} \in \operatorname{col}(A) \textit{ pour tout } \mathbf{b} \in \mathbb{R}^m \\
&\quad \iff \mathbb{R}^m \textit{ est engendré par les colonnes de } A \\
&\quad \iff \mathbb{R}^m = \operatorname{col}(A) \\
&\quad \iff \textit{aucune rangée contrainte dans la forme échelonnée de } [A|b] \\
&\quad \iff \textit{aucune rangée nulle dans la forme échelonnée de } A \\
&\quad \iff \operatorname{rang}(A) = m
\end{aligned}$$

On peut alors déterminer un ensemble engendrant pour $\mathbb{R}^m$: on retient les colonnes qui correspondent aux pivots! Ce serait un ensemble engendrant minimal, car si on enlève un, on aurait alors une rangée nulle dans la forme échelonnée de A.

On peut combiner les deux résultats.

Théorème 9.

$$\begin{aligned}
&A\mathbf{x} = \mathbf{b} \textit{ est consistant pour tout } \mathbf{b} \in \mathbb{R}^m \textit{ et } A\mathbf{x} = \mathbf{0} \textit{ implique } \mathbf{x} = \mathbf{0} \\
&\quad \iff \textit{les colonnes de } A \textit{ sont indépendants et engendrent } \mathbb{R}^m \\
&\quad \iff \textit{les colonnes de } A \textit{ sont une base pour } \mathbb{R}^m \\
&\quad \iff \operatorname{rang}(A) = m \textit{ et } \operatorname{rang}(A) = n \textit{ (et donc } m = n\textit{)} \\
&\quad \iff \textit{chaque colonne et chaque rangée de } A \textit{ possède un pivot} \\
&\quad \iff A\mathbf{x} = \mathbf{b} \textit{ possède une solution unique pour tout } \mathbf{b} \in \mathbb{R}^m
\end{aligned}$$

1. Une Forme Échelonnée et Réduite

Soit A une matrice: il y a quatre espaces qui en découlent: $\mathrm{col}(A)$, $\ker(A)$, $\mathrm{col}(A^T)$ et $\ker(A^T)$. On cherche une méthode pour déterminer ces espaces. Comme d'habitude on cherche une méthode *paresseuse*, c'est-à-dire on veut autant que possible rétuliser nos calculs.

Exemple: Voici une matrice A, ainsi que sa forme échelonnée et échelonnée réduite, et aussi pour la matrice augmentée avec une colonne de constantes arbitraires. Comme exercice, vérifier!

$$A = \begin{bmatrix} -1 & 2 & -2 & 0 & 6 & -1 \\ 0 & 0 & 1 & 1 & -2 & 1 \\ 1 & -2 & 1 & -1 & -4 & 1 \\ -1 & 2 & -4 & -2 & 10 & -2 \\ 3 & -6 & 5 & -1 & -16 & 4 \end{bmatrix} \to \cdots \begin{bmatrix} 1 & -2 & 2 & 0 & -6 & 1 \\ 0 & 0 & 1 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\to \cdots \begin{bmatrix} 1 & -2 & 0 & -2 & -2 & 0 \\ 0 & 0 & 1 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$[A|\mathbf{b}] = \left[\begin{array}{cccccc|c} -1 & 2 & -2 & 0 & 6 & -1 & a \\ 0 & 0 & 1 & 1 & -2 & 1 & b \\ 1 & -2 & 1 & -1 & -4 & 1 & c \\ -1 & 2 & -4 & -2 & 10 & -2 & d \\ 3 & -6 & 5 & -1 & -16 & 4 & e \end{array}\right] \to \cdots \left[\begin{array}{cccccc|c} 1 & -2 & 2 & 0 & -6 & 1 & a \\ 0 & 0 & 1 & 1 & -2 & 1 & b \\ 0 & 0 & 0 & 0 & 0 & 1 & a+b+c \\ 0 & 0 & 0 & 0 & 0 & 0 & -2a+b-c+d \\ 0 & 0 & 0 & 0 & 0 & 0 & a-b-2c+e \end{array}\right]$$

$$\to \cdots \left[\begin{array}{cccccc|c} 1 & -2 & 0 & 0 & -6 & 0 & -2a-b-c \\ 0 & 0 & 1 & 1 & -2 & 0 & -a-c \\ 0 & 0 & 0 & 0 & 0 & 1 & a+b+c \\ 0 & 0 & 0 & 0 & 0 & 0 & -2a+b-c+d \\ 0 & 0 & 0 & 0 & 0 & 0 & a-b-2c+e \end{array}\right]$$

On cherche à savoir quelles formes sont *nécessaires*. On ne veut pas tout calculer si on peut s'en débrouiller avec une seule!

2. Bases Pour col(A)

On cherche l'espace colonne: précisément, une base. Rapelons que

$$\mathrm{col}(A) = \{\mathbf{b} \mid A\mathbf{x} = \mathbf{b} \text{ est consistant } \}$$

Donc on cherche les vecteurs $\mathbf{b}$ tel que $[A|\mathbf{b}]$ ne possède aucune rangée contradictoire. La troisième forme se montre utile: on voit qu'il n'y aurait aucune rangée contradictoire exactement lorsque:

$$0 = -2a + b - c + d \qquad 0 = a - b - 2c + e$$

Donc

$$\begin{aligned}\mathrm{col}(A) &= \left\{ \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} \;\middle|\; \begin{array}{l} 0 = -2a + b - c + d \\ 0 = a - b - 2c + e \end{array} \right\} \\ &= \left\{ \begin{bmatrix} a \\ b \\ c \\ 2a - b + c \\ -a + b + 2c \end{bmatrix} \right\} = \left\{ a \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \\ -1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} \right\}\end{aligned}$$

On voit alors que

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} \right\}$$

est un ensemble engendrant. C'est aussi un ensemble indépendant, à cause des trois premières positions (mettre une combinaison linéaire égale à $\mathbf{0}$, et regarder les trois premières équations. Donc c'est une base. Note que ces trois position correspondent exactement aux pivots: ce serait toujours une base!

On n'a rien dû calculer de plus. Par contre, on a utilisé une forme ayant une colonne de constantes. On préfère une méthode qui utilise la forme échelonnée de A seul.

On connait maintenant $\dim \mathrm{col}(A) = 3$. Pourquoi 3? On avait 5 inconnues a, b, c, d, e. En substituant les deux conditions, on a perdu deux. En fait on a perdu exactement les deux inconnues qui correspondent aux colonnes sans pivots. Donc la dimension est exactement le nombre de colonnes avec pivot.

Soit B la sous-matrice de A correspondant aux colonnes ayant des pivvots. Quelle est la forme échelonnée de B? C'est exactement la sous matrice de la forme échelonnée de A! Bref: la forme échelonnée de B possède le même nombre de pivots que A — en fait, les mêmes pivots! Alors B possède un pivot par colonne, donc les colonnes de B sont indépendants.

On connaît alors que si on retient les colonnes de A avec pivot, on obtient un ensemble de trois vecteurs dans un espace de dimension 3 (i.e., l'espace $\mathrm{col}(A)$!). Donc, c'est une base!

Théorème 1. *Afin de trouver une base pour* $\mathrm{col}(A)$ *on a deux options:*

- *Trouver une forme échelonnée pour la matrice augmentée* $[A\mathbf{b}]$. *Les rangées nulle (à la gauche de la barre) donnent les conditions, qu'on substitue dans* $\mathbf{b}$. *En écrivant* $\mathbf{b}$ *en termes de chaque inconnue, on découvre une base.*
- *Trouver une forme échelonnée pour A. Retenir les colonnes de A ayant pivot.*

On a $\dim \mathrm{col}(A) = \mathrm{rang}(A)$.

À vous de décider quelle méthode est plus simple! Mais attention: dans la deuxième, il faut absolument prendre les colonnes de A et non de la forme échelonnée. Pourquoi?

NB: La première méthode donne les conditions pour l'espace colonne aisni qu'une base. La deuxième donne seulement une base, mais c'est plus directe. On verra plus tard que la première est utile pour d'autres choses.

3. Bases Pour $\ker(A)$

On se rapelle la définition:

$$\ker(A) = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{0}\}$$

Il s'agit simplement de solutionner le système homogène. Puisqu'il y aura toujours des zéros à la droite de la barre, on peut l'omettre. La forme échelonnée réduite de A est utile. Note que la première rangée de la forme échelonnée semble dit

$$x_1 - 2x_2 + 2x_3 - 6x_5 + x_6 = 0$$

Ou est le "$= 0$"? C'est la colonne des constantes, tous zéros, qu'on n'a pas écrit. Sachant ceci, on lit la solution générale directement: c'est exactement $\ker(A)$!

$$\ker(A) = \left\{ \begin{bmatrix} 2x_2 + 2x_4 + 2x_5 \\ x_2 \\ -x_4 + 2x_5 \\ x_4 \\ x_5 \\ 0 \end{bmatrix} \right\} = \left\{ x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \mid x_2, x_4, x_5 \in \mathbb{R} \right\}$$

On a donc un ensemble engendrant pour $\ker(A)$:

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

C'est aussi un ensemble indépendant, à cause des positions 2, 4, et 5. Encore une fois, mettre une combinaison linéaire égale à zéro: ces équations forcent la solution triviale. Elles correspondent exactement aux paramètres. Donc on aurait toujours une base.

> **Théorème 2.** *Afin de trouver une base pour* $\operatorname{col}(A)$ *on trouve la forme échelonnée réduite pour* A*, et on donne alors la solution générale. En séparant les paramètres, on obtient une base.*
>
> *On a* $\dim \ker(A) = n - \operatorname{rang}(A)$

4. Bases Pour $\operatorname{col}(A^T)$

On connait comment trouver une base pour l'espace colonne. Donc en principe, il suffit de poser $B = A^T$ et refaire la méthode de Gauss-Jordan sur B. C'est correcte, mais on veut éviter de faire d'autre travail.

La clé c'est que les opérations de rangées *ne changent pas l'espace engendré par les rangées*. Donc les rangées de la forme échelonnée réduite engendrent aussi $\operatorname{col}(A^T)$ (ainsi que les rangées à n'importe quelle étape intermédiare). On cherche un ensemble engendrant minimal. Ceci se voit plus facilement dans la forme échelonnée réduite: c'est exactement les rangées non-nulles. Les rangées avec pivot seront toujours indépendants (mettre une combinaison linéaire égal à zéro, et regarder le système linéaire qui en découle). Certainement les rangées nulles ne seront jamais dans une base (pourquoi?). Donc les rangées avec pivots forment une base pour $\operatorname{col}(A^T)$. La dimension est exactement le nombre de pivots.

Théorème 3. *Afin de trouver une base pour* $\operatorname{col}(A^T)$*, on prend les rangées avec pivots à n'importe quelle moment de la réduction.*

On a $\dim \operatorname{col}(A^T) = \operatorname{rang}(A)$.

Note qu'on n'a pas besoin de la forme échelonnée réduite: il suffit de connaître les pivots.

Note que en principe on *pourrait* poser $B = A^T$, et trouver une base pour $\operatorname{col}(A^T)$ en trouvant une base pour $\operatorname{col}(B)$. On aura donc que $\dim \operatorname{col}(A^T) = \dim \operatorname{col}(B) = \operatorname{rang}(B)$. Mais on conaît que $\dim \operatorname{col}(A^T) = \operatorname{rang}(A)$. Puisque $B = A^T$ on a:

Théorème 4. $\operatorname{rang}(A) = \operatorname{rang}(A^T)$

Ce n'est *pas* vrai que les pivots de A sont les pivots de A^T. Ce n'est pas non plus vrai que la forme échelonnée de A^T est la transpose de la forme échelonnée de A.

Par contre le *nombre* de pivots est invariante! Le rang d'une matrice ne change pas si on fait des opérations de rangée, ou même des transposes.

5. Bases Pour $\ker(A^T)$

En principe, on pourrait poser $B = A^T$ et refaire la méthode de Gauss-Jordan pour trouver $\ker(B)$. On ne fera pas ceci. Pour le moment on ne fera rien! Mais on verra une méthode élégante. . . qui est relié à la première méthode pour $\operatorname{col}(A)$.

Note qu'en pensant à faire $\ker(B)$ avec A^T on obient au moins que la dimension est le nombre de rangées sans pivot.

Théorème 5. $\dim \ker(A^T) = m - \operatorname{rang}(A)$

6. Lire une Matrice

Donc, pour trouver une base pour l'espace colonne et l'espace rangée, il suffit de connaître les pivots. Pour trouver une base pour le noyau, il faut trouver la forme échelonné réduite (ou solutionner le système homogène).

Exemple: Donner une base pour $\ker(B)$, avec $B = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 2 & 4 \end{bmatrix}$.

On voit que

$$\begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 2 & 4 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Il y a alors un pivot dans chaque colonne. Les trois rangées disent:

$$x_1 = 0 \qquad x_2 = 0 \qquad 0 = 0$$

La solution générale homogène (i.e., le noyau) est:

$$\ker(B) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

Ou est la base? On ne peut pas prendre le vecteur $\mathbf{0}$ pour une base (pourquoi?). La réponse c'est que la base est vide: $\emptyset$. Vue d'une autre manière la dimension du noyau est

$$\dim \ker(B) = n - \operatorname{rang}(B) = 2 - 2 = 0$$

Donc la taille d'une base est zéro!

En passant, on voit que $\dim \operatorname{col}(B) = \dim \operatorname{col}(B^T) = \operatorname{rang}(B) = 2$ et $\dim \ker(B^T) = m - \operatorname{rang}(B) = 3 - 2 = 1$.

Pourquoi pas continuer? Une base pour $\operatorname{col}(B)$ est $\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} \right\}$. Une base pour $\operatorname{col}(B^T)$ est $\left\{ \begin{bmatrix} 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix} \right\}$.

De plus, grâce au pivots, on connaît que les rangées de B sont dépendants et les colonnes sont indépendantes. Si B est la matrice des coefficients d'un système linéaire il y aura soit aucune solution soit une solution unique. Donc si le système est consistant il y aura une solution unique. Les rangées engendrent $\mathbb{R}^n = \mathbb{R}^2$ mais les colonnes n'engendrent pas $\mathbb{R}^m = \mathbb{R}^3$.

Ce sont des conséquences des Quatre Théorèmes.

Exercice: Soit $C = \begin{bmatrix} 0 & 3 & 1 & 4 \\ 0 & 2 & -1 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \to \cdots \to \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Donner des bases pour $\operatorname{col}(C)$, $\operatorname{col}(C^T)$ et $\ker(C)$. Donner les dimensions des quatre espaces. Dire si les rangées et colonnes de C sont indépendants (avec référence à la forme échelonnée réduite. Expliquer combien de solutions le système linéaire pourrait avoir. Qu'est-ce qui engendre quoi? Qui est indépendant? Et…? Est-ce que c'était vraiment nécessaire de trouver la forme échelonnée réduite? Quand?

1. Annuler

On a déja vu des exemples de matrices A, X, Y tel que $AX = AY$ mais $X \neq Y$. Voici une autre :

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 2 & 5 \end{bmatrix}$$

Par contre, parfois si $AX = AY$, on peut avoir $X = Y$ aussi. Voici un exemple. Imaginons que $AX = AY$ et que $A = I$. Alors $AX = IX = X$ et $AY = IY = Y$. AX et AY étant égaux, on a $X = Y$. Ce n'est pas vraiment un exemple intéressant (c'est comme dire « Si $1x = 1y$ alors $x = y$ » pour des chiffres x et y). La question intéressante est :

Quelles conditions sur A garantissent que si $AX = AY$, on a $X = Y$?

Si $AX = AY$, alors $A(X - Y) = \mathbf{0}$.[1] L'équation $A(X - Y) = \mathbf{0}$ se comprend plus facilement avec un dessin montrant les colonnes de la matrice $X-Y$: chaque colonne de $X - Y$ est dans le noyau de A. On veut une condition sur A qui *oblige* $X = Y$; donc on veut une condition sur A qui *oblige* chaque colonne de $X - Y$ à être $\mathbf{0}$.[2] Évidemment, il faut exiger que $\ker(A) = \{\mathbf{0}\}$!

Si $\ker(A) = \{\mathbf{0}\}$, alors les colonnes de $X - Y$ seront tous zéros, et donc $X = Y$. Par contre si $\ker(A) \neq \{\mathbf{0}\}$, il existe un vecteur non-nulle $\mathbf{d} \in \ker(A)$. Posons D comme une matrice ayant chaque colonne égale à $\mathbf{d}$, et $Y = X + D$. Peu importe X, on aura toujours $X \neq Y$; de plus on aura toujours $AY = A(X + D) = AX + AD = AX = \mathbf{0} = AX$.

On a prouvé que :

Théorème 1. *On a toujours le droit d'annuler le A dans $AX = AY$ si et seulement si* $\ker(A) = \{\mathbf{0}\}$.[3]

[1] C'est la matrice $\mathbf{0}$. C'est quoi la relation entre les tailles de A, X, Y et $\mathbf{0}$?

[2] C'est le vecteur $\mathbf{0}$. C'est quoi la taille de *ce* $\mathbf{0}$?

[3] si et seulement si les colonnes de A sont indépendantes si et seuelement si il y a un pivot dans chaque colonne. . .

Posons $B = A^T$. Alors on connait que les rangées de B sont indépendants, et que B possède un pivot dans chaque rangée. Donc $B\mathbf{y} = \mathbf{b}$ est consistant pour tout $\mathbf{b}$.[4] On peut donc choisir $\mathbf{b}$ n'importe comment, sachant qu'il existerait une solution. Par exemple, il existe un vecteur z_1 tel que $B\mathbf{z}_1$ donne la première colonne de I. Il existe aussi des vecteurs $\mathbf{z}_2, \ldots, \mathbf{z}_m$ tel que $B\mathbf{z}_i$ donne la i-ième colonne de I. Si on met ces vecteurs $\mathbf{z}_1, \ldots, \mathbf{z}_m$ comme les colonnes d'une matrice Z on voit que $BZ = I$. Donc il existe une matrice Z tel que :

$$\begin{aligned} & BZ = I \\ \iff & (BZ)^T = I^T \\ \iff & Z^T B^T = I^T \\ \iff & Z^T A = I \end{aligned}$$

Donc si $\ker(A) = (\mathbf{0})$, il existe une matrice Z^T tel que $Z^T A = I$, qui permet de faire :

$$\begin{aligned} & AX = AY \\ \Longrightarrow & Z^T AX = Z^T AY \\ \Longrightarrow & IX = IY \\ \Longrightarrow & X = Y \end{aligned}$$

On dit que Z^T est une INVERSE À GAUCHE pour A.

Théorème 2. *A possède une inverse à gauche si et seulement si* $\ker(A) = \{\mathbf{0}\}$. *Dans ce cas $AX = AY$ implique que $X = Y$.*

On peut remplacer la condition « $\ker(A) = (\mathbf{0})$ » par l'équivalente « les colonnes de A sont indépendantes » ou « $\text{rang}(A) = n$ ».

D'une manière similaire, on à que

Théorème 3. *A possède une inverse à droite si et seulement si* $\ker(A^T) = \{\mathbf{0}\}$. *Dans ce cas $XA = YA$ implique que $X = Y$.*

On peut remplacer la condition « $\ker(A^T) = (\mathbf{0})$ » par l'équivalente « les rangées de A sont indépendantes » ou « $\text{rang}(A) = m$ »ou ...

On résume :

[4] Et les tailles de $\mathbf{y}$ et $\mathbf{b}$ en termes de A sont... ?

Théorème 4. *Soit A de taille $m \times n$.*

A possède une inverse à gauche $\iff \operatorname{rang}(A) = n$
A possède une inverse à droite $\iff \operatorname{rang}(A) = m$
A possède une inverse à gauche et à droite $\iff \operatorname{rang}(A) = n = m$
A possède ni une inverse à gauche ni à droite $\iff \operatorname{rang}(A) < m, n$

Exercice: Revoir les Quatre Théoèmes, et ajouter des formes équivalentes pour les inverses ! Donner une liste de choses qui sont équivalentes à « A possède une inverse à gauche / à droite / les deux / aucun ».

2. Carré

On a déja vu qu le cas $\operatorname{rang}(A) = m = n$ est importante : ça correspond au cas ou tout système $A\mathbf{x} = \mathbf{b}$ possède une solution unique. On connaît maintenant de plus que ça veut dire que A possède une inverse gauche et une inverse à droite. On s'intéresse au cas spécial où l'inverse à gauche et l'inverse à droite sont identiques.

Définition 5. *Soit A et B des matrices de taille $n \times n$.*

Si $AB = BA = I$ alors on dit que B est une inverse de A. On écrit que $A^{-1} = B$. On dit que A est inversible.

La première chose à voir c'est que B est *la* inverse.

Théorème 6. *Si A possède une inverse, alors cette inverse est unique.*

La démonstration est directe.[5] Supposons que B_1 et B_2 sont des inverses de A. Alors

$$B_1 = B_1 (AB_2) = B_1 AB_2 = (B_1 A) B_2 = B_2$$

Donc $B_1 = B_2$, et A ne possède qu'une seule inverse. $\square$

Le mot « unique » est relié aux inverses d'une deuxième manière :

Théorème 7. *Si A possède une inverse, alors pour tout $\mathbf{b}$ le système $A\mathbf{x} = \mathbf{b}$ possède une solution unique.*

[5] La comparer avec le démonstration que dans un espace vectoriel le vecteur zéro est unique, et l'inverse d'un vecteur est unique

On connaît déja ce théorème. Si vous avez fait l'exercice précédent, vous l'avez déja démontré ! Mais voici une *autre* preuve. Si A possède une inverse alors on peut faire :

$$\begin{aligned} &A\mathbf{x} = \mathbf{b} \\ \Longrightarrow &A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \\ \Longrightarrow &I\mathbf{x} = A^{-1}\mathbf{b} \\ \Longrightarrow &\mathbf{x} = A^{-1}\mathbf{b} \end{aligned}$$

Donc la (oui, *la*) solution est $\mathbf{x} = A^{-1}\mathbf{b}$. C'est utile pour solutionner des systèmes linéaires.

Exemple: Solutionner $\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

On vous dit que $\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$. Vous êtes très sceptique. D'où vient cette matrice inverse ? Vous décidez toute de suite de prendre un papier brouillon afin de vérifier que vraiment le produit donne I, dans les deux sens !

De retour, on peut maintenant donner la solution :

$$\mathbf{x} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Vous posez certainement deux questions : d'où vient cette matrice, et pourquoi solutionner le système comme ceci : on vous a dit que la méthode de Gauss-Jordan était optimale ?

On répondra à la première question dans quelques instants, mais pour la deuxième, on note que dans les applications, c'est souvent le cas qu'il faut solutionner un système "variable", c'est-à-dire qu'on veut re-solutionner avec un autre $\mathbf{b}$.

Exemple: Solutionner $\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$

On connaît déja l'inverse. On peut écrire la solution tout de suite :

$$\mathbf{x} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 16 \\ -4 \end{bmatrix}$$

Dans le cas d'un système plus grand, c'est très utile de pouvoir réutiliser le travail qu'on a fait pour déterminer l'inverse de A.

3. Comment

Comment trouver l'inverse de A ?

Commençons par un exemple. On cherche l'inverse de $\begin{bmatrix}1 & 2\\ 1 & 3\end{bmatrix}$, c'est-à-dire on cherche une matrice $\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ tel que $\begin{bmatrix}1 & 2\\ 1 & 3\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$.

Donc on veut résoudre les deux systèmes :

$$\left[\begin{array}{cc|c} 1 & 2 & a\\ 1 & 3 & c\end{array}\right] \qquad \left[\begin{array}{cc|c} 1 & 2 & b\\ 1 & 3 & d\end{array}\right]$$

On voit déja un avantage : en trouvant une forme échelonnée on verra le nomber de pivots : en particulier, on saura si l'inverse existe ou non. L'inverse existe exactement lorsqu'il y a un pivot dans chaque colonne de A. Puisque A est carrée on voit que :

Théorème 8. *A possède une inverse si et seulement si* $\operatorname{rang}(A) = m = n$ *si et seuelement si la forme échelonnée réduite de A est la matrice I.*

Si la forme échelonnée est I, la solution est particulièrement simple à voir.

(**Exercice:** Donner la solution générale de $\left[\begin{array}{cc|c} 1 & 0 & 13\\ 0 & 1 & -14\end{array}\right]$)

On veut solutionner deux systèmes linéaires ; ce n'est pas un énorme travail, mais en générale ce serait autant de systèmes que A possède colonnes. On cherche à économiser. On les fera tous ensemble !

$$\xrightarrow[R_2 \to R_2 - R_1]{} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0\\ 0 & 1 & -1 & 1\end{array}\right] \xrightarrow[R_1 \to R_1 - 2R_2]{} \left[\begin{array}{cc|cc} 1 & 0 & 3 & -2\\ 0 & 1 & -1 & 1\end{array}\right]$$

Comment lire cette matrice « sur-augmentée » ? En regardant à la première colonne d̀roite de la barre, on voit la solution au premier système : $x_1 = 3$ et $x_2 = -1$. C'est donc la première colonne de notre inverse. En regardant à la deuxième colonne à droite de la barre, on voit la solution au deuxième système : $x_1 = -2$ et $x_2 = 1$. C'est la deuxième colonne de notre inverse. Donc notre inverse est exactement :

$$\begin{bmatrix}3 & -2\\ -1 & 1\end{bmatrix}$$

C'est exactement la matrice à la droite de la barre !

Exercice: Vérifier que $\begin{bmatrix}1 & 2\\ 1 & 3\end{bmatrix}\begin{bmatrix}3 & -2\\ -1 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ et $\begin{bmatrix}3 & -2\\ -1 & 1\end{bmatrix}\begin{bmatrix}1 & 2\\ 1 & 3\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$

Algorithme 9. *Comment trouver A^{-1}.*

(1) *Former la matrice augmentée $[A|I]$.*
(2) *trouver la forme échelonnée réduite.*
(3) *Si* $\operatorname{rang}(A) < n$ *alors il n'existe aucune inverse.*

> *(4) Si* $\operatorname{rang}(A) = n$ *alors on a* $[I|A^{-1}]$ *: l'identité à la gauche et l'inverse à la droite.*

Note que pour solutionner un système d'équations selon la méthode de Gauss-Jordan on réduit une matrice de taille $m \times (n+1)$. Pour la solutionner selon la méthode des inverses on réduit une matrice de taille $m \times (2n)$. C'est plus de travail, mais si on a plusieurs systèmes, on épargne au long.

Il reste un détail. On a trouvé une matrice B tel que $AB = I$. Est-ce vraiment l'inverse ? Comment sait-on que en générale $BA = I$?

> **Théorème 10.** *Soit A et B des matrices de taille $n \times n$.*
>
> *Si $AB = I$ alors $BA = I$ aussi.*

Supposons que A et B sont $n \times n$ et que $AB = I$. Donc A possède certainement une inverse à droite : B. Donc A possède un pivot dans chaque rangée, et donc (puisque A est carré) un pivot dans chaque colonne. Donc A possède une inverse à gauche. Posons alors C l'inverse à gauche ; donc $CA = I$. Alors :

$$C = CI = C(AB) = CAB = (CA)B = IB = B$$

Donc $C = B$.

On voit que si A possède une inverse à gauche et une inverse à droite, c'est la même inverse.

Exemple: Trouver l'inverse de $P = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ -1 & 0 & 2 \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3 \to R_3 + R_1]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3 \to -R_3]{} \left[\begin{array}{ccc|ccc} 1 & 0 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{array}\right]$$

$$\xrightarrow[R_1 \to R_1 + 3R_3]{R_2 \to R_2 - 2R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & -3 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{array}\right]$$

On a $P^{-1} = \begin{bmatrix} -2 & 0 & -3 \\ 2 & 1 & 2 \\ -1 & 0 & -1 \end{bmatrix}$.

Exemple: Trouver l'inverse de $Q = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ -1 & -3 & 1 \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ -1 & -3 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[R_3 \to R_3 + R_1]{} \left[\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -1 & -2 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3 \to R_3 + R_2]{} \left[\begin{array}{ccc|ccc} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

On voit que $\text{rang}(Q) = 2 < n$. Donc A n'est pas inversible.

Exemple: Trouver l'inverse de $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Pour que l'inverse existe, il faut que $\text{rang}(A) = 2$. De façon équivalente, il faut que les colonnes de A sont indépendantes. Ceci veut dire qu'elles ne sont pas multiples l'un l'autre. Si elles sont multiples l'un l'autre, alors $b = ta$ et $d = tc$, pour $t \neq 0$. Donc $tcb = tad$. C'est possible que a, b, c ou d soient zéro, mais pas t. Donc on peut diviser par t pour obtenir que $ad - bc = 0$.

Alors A est inversible exactement lorsque $ad - bc \neq 0$.

On a soit $a = 0$ et $c \neq 0$, soit $a \neq 0$ et $c = 0$, soit $a \neq 0$ et $c \neq 0$.

On peut faire la réduction de Gauss-Jordan pour chaque cas, afin de donner une formule pour l'inverse. On constate alors que à la droite de la barre on trouve :

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Note que cette formule est invalide quand $ad - bc = 0$, car on ne peut pas diviser par zéro. Mais si $ad - bc = 0$ alors il n'y a pas d'inverse, donc il ne peut exister de formule.

4. Défi

On connaît maintenant comment trouver l'inverse d'une matrice carré. Peut-on adapter cette méthode pour trouver les inverses à gauche et à droite ?

La réponse est oui, mais il y a un problème. Si C est tel que $n = \text{rang}(C)$, alors on connît que C possède une inverse à gauche. Mais ce ne serait pas unique ; il y aurait un nombre infini !

Donc, pour ceux qui veulent un défi :

Expliquer *pourquoi* l'inverse à gauche n'est unique que si C est carré. Trouver une algorithme pour déterminer alors *toutes* les inverses à gauche de C. L'algorithme devrait donner trois réponses possibles : soit que C ne possède aucune inverse à gauche, soit donner l'inverse à gauche unique de C (dans tel cas C est obligatoirement carré), soit donner toutes les inverses à gauche de C.

Note que si vous êtes capable de faire les inverses à gauche, l'inverse à droite de C est la même chose que l'inverse à gauche de C^T...

5. Arithmétique

Voici quelques propriétés des inverses qui sont souvent utiles.

(1) $I^{-1} = I$.

(2) Si A est inversible alors A^{-1} possède une inverse et $\left(A^{-1}\right)^{-1} = A$.

(3) Si A et B sont des matrices inversibles de taille $n \times n$, alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.

(4) Si A est inversible alors A^k est inversible pour tout k et $\left(A^k\right)^{-1} = \left(A^{-1}\right)^k$.

(5) Si A est inversible alors cA est inversible pour tout $c \neq 0$ et $(cA)^{-1} = \frac{1}{c}A^{-1}$.

Note que l'inverse renverse l'ordre d'un produit : c'est un peu comme les transposes.

Exemple: Simplifier $A^3(BA^2)^{-1}$

$$\begin{aligned} A^3(BA^2)^{-1} &= A^3 \left(A^2\right)^{-1} B^{-1} \\ &= A^3 \left(A^{-1}\right)^2 B^{-1} \\ &= (A)(A)(A) \left(A^{-1}\right) \left(A^{-1}\right) B^{-1} \\ &= (A)(A) \left(A^{-1}\right) B^{-1} \\ &= AB^{-1} \end{aligned}$$

Note que $A^3 \left(A^{-1}\right)^2 = A = A^{3-2}$. C'est très générale. On écrit souvent A^{-2} pour $\left(A^{-1}\right)^2$. Donc $A^1 = 1$ et $A^0 = I$.

Exemple: Simplifier $(A + B)^{-1}$.

On ne peut rien faire. Ce n'est certainement pas vrai que c'est $A^{-1} + B^{-1}$. Même si A et B sont inversibles, $A + B$ pourrait ne pas l'être.

Exemple: Simplifier $B\left(A^2B\right)^{-1} A^3$.

$$B\left(A^2B\right)^{-1} A^3 = BB^{-1}\left(A^2\right)^{-1} A^3 = BB^{-1}\left(A^{-1}\right)^2 A^2A = A$$

1. Complément Orthogonal d'un Espace

Rappelons deux définitions. L'espace rangée est l'espace de toutes les combinaisons linéaires des rangées de A (ou des colonnes de A^T. Le noyau est l'espace de touts les vecteurs qui sont "annulés" par A. En symboles :

$$\operatorname{col}(A^T) = \left\{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} = A^T\mathbf{z} \text{ pour un vecteur } \mathbf{z}\right\}$$
$$\ker(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$$

La multiplication matricielle de A et $\mathbf{x}$ consiste en prendre le produit scalaire de chaque rangée de A avec $\mathbf{x}$. Aussi, deux vecteurs sont orthogonaux si et seulement si leur produit scalaire est zéro. Donc $\mathbf{x} \in \ker(A)$ si et seulement si $\mathbf{x}$ est orthogonal à chaque rangée de A. Être orthogonal à chaque rangée de A est équivalant à être orthongal à toute combinaison linéaire des rangées ; bref, être orthogonal à l'espace $\operatorname{col}(A^T)$.

Pour ceux qui préfèrent une preuve en symboles, prenons $\mathbf{x} \in \ker(A)$ et $\mathbf{y} \in \operatorname{col}(A^T)$. Alors

$$\mathbf{y} \cdot \mathbf{x} = \mathbf{y}^T\mathbf{x} = \left(A^T\mathbf{z}\right)^T\mathbf{x} = \mathbf{z}^TA\mathbf{x} = \mathbf{z}^T\mathbf{0} = 0$$

Donc on a trouvé que tout vecteur dans $\ker(A)$ est orthogonal à tout vecteur dans $\operatorname{col}(A^T)$. C'est un concept suffisament important pour mériter un nom en général :

> **Définition 1.** *Soit U un sous-espace de $\mathbb{R}^n$.*
>
> *Alors $U^\perp$ est l'ensemble des vecteurs qui sont orthogonaux à tout vecteur de U. On dit que $U^\perp$ est le* COMPLÉMENT ORTHOGONAL *de U dans $\mathbb{R}^n$.*

Exemple: Soit l'espace $P = \left\{\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + 3z = 0\right\}$.

C'est un plan dans $\mathbb{R}^2$. Une base pour P est $\left\{\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}\right\}$. Un vecteur normal est $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ (ou un multiple de ce vecteur). L'ensemble $P^\perp$ est l'ensemble de tout vecteur orthogonal au plan. Un vecteur qui est orthongal au plan est soit zéro, soit un vecteur normal du plan. Donc $P^{\perp} = \left\{ t \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \mid t \in \mathbb{R} \right\}$. Les vecteurs *dans* le plan forment P. Les vecteurs *normal* au plan forment $P^{\perp}$.

On voit dans cet exemple que $P^{\perp}$ est un *espace vectoriel*. C'est complètement général.

Théorème 2. *Soit U sous-espace de $\mathbb{R}^n$. Alors $U^{\perp}$ est sous-espace de $\mathbb{R}^n$.*

Exercice: Démontrer ce théorème, à l'aide du test de sous-espace. C'est-à-dire, montrer que (i) $\mathbf{0} \in U^{\perp}$; (ii) si $\mathbf{y}, \mathbf{z} \in U$, alors $\mathbf{y} + \mathbf{z} \in U$; (iii) si $\mathbf{y} \in U$ alors $k\mathbf{y} \in U$.

On connaît que $P^{\perp}$ est l'espace de touts les vecteurs normaux. Quels vecteurs sont orthogonaux à touts les vecteurs normauxde P ? Un dessin montre que c'est exactement les vecteurs dans le plan original P. Donc $\left(P^{\perp}\right)^{\perp} = P$.

Quels vecteurs sont à la fois dans P et dans $P\perp$? Un tel vecteur devrait être orthogonal à soi-même, donc le vecteur $\mathbf{0}$: alors $P \cap P^{\perp} = \{\mathbf{0}\}$

On voit aussi que $\dim(P) + \dim(P^{\perp}) = 3$.

Théorème 3. *Soit U un sous-espace de $\mathbb{R}^n$. Alors*
(1) $\left(U^{\perp}\right)^{\perp} = U$
(2) $U \cap U^{\perp} = \{\mathbf{0}\}$
(3) $\dim(U) + \dim(U^{\perp}) = n$

La notation « $U \cap U^{\perp}$ » veut dire « touts les vecteurs qui sont dans U et dans $U^{\perp}$ ». Donc :

Le seul vecteur qui est dans U et dans $U^{\perp}$ est le vecteur $\mathbf{0}$.

On verra une preuve de la dernière partie du théorème dans un instant.

2. Espaces Orthogonaux et Matrices I

Nos observations de la section précédente montrent exactement que pour une matrice A on a

Théorème 4. $\ker(A) = \text{col}(A^T)^\perp$

De plus, soit U n'importe quel sous-espace de $\mathbb{R}^n$. U possède une base, et donc on peut former une matrice A tel que les rangées de A forment une base pour U. Alors

$$\dim(U) + \dim(U^\perp) = \dim(\text{col}(A^T)) + \dim(\ker(A)) = n$$

Exemple: Trouver une base pour $\ker(A)$, où $A = \begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$.

$$\begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -\frac{1}{3} \\ 0 & 1 & -\frac{2}{3} \end{bmatrix}$$

On lit la solution générale au système homogène :

$$\ker(A) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mid x_1 = \frac{1}{3}x_3, x_2 = \frac{2}{3}x_3 \right\} = \left\{ t \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ 1 \end{bmatrix} \right\} = \left\{ t \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

Donc une base pour $\ker(A)$ est $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. On voit de plus que les rangées de A forment une base pour l'espace P de la première section, car il y a un pivot dans chaque rangée.

Que montre vraiment cet exemple ? Sachant un ensemble de vecteurs qui engendrent un espace P, on forme une matrice A avec ces vecteurs comme rangées. En trouvant la forme échelonnée réduite de A on peut déduire une base pour P et une base pour $P^\perp$.

Exemple: Soit les deux vecteurs $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$ dans $\mathbb{R}^3$. Trouver un vecteur orthogonal à ces deux vecteurs.

C'est exactement le précédent...

Algorithme 5. *Soit $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ des vecteurs dans $\mathbb{R}^n$, et U l'espace qu'ils engendrent. Afin de trouver une base pour $U^\perp$:*

(1) On met les vecteurs en rangée dans une matrice A.

(2) On réduit la matrice à la forme échelonnée réduite afin de trouver $\ker(A)$.

(3) On substitue les conditions pour obtenir une base pour $\ker(A)$.

On comprend ceci en termes de la géométrie. La base pour U donne les « directions permises » dans U et la base pour $U^\perp$ donne les « directions défendues ». Puisque

$U = \mathrm{col}(A^T)$ et $U^\perp = \ker(A)$, on peut trouver ces bases par des méthodes purement matricielles qu'on connaît déja !

Pour un plan il y a deux directions permises (ou vecteurs directeurs) et un vecteur normal. En général, il peut en avoir plus.

3. Un Exemple I

Exemple: Soit H l'espace engendré par $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 5 \\ 1 \\ 2 \end{bmatrix} \right\}$

Trouver une base pour H, et une base pour $H^\perp$.

On met les vecteurs en rangée, et on réduit :

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 3 & 1 & 0 \\ 1 & 1 & 5 & 1 & 2 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Il y a un pivot dans chaque rangée. Donc les trois rangées originaux[1] forment une base pour $\mathrm{col}(A^T)$. Donc les trois vecteurs originaux forment une base pour H. Si on avait eu deux pivots, on aurait retenu seulement deux vecteurs pour former la base (et alors H serait un plan en $\mathbb{R}^5$)

On lit la solution au système homogène pour donner une base pour $\ker(A)$. On obtient la base :

$$\left\{ \begin{bmatrix} -2 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \right\}$$

Ce sont *deux* vecteurs normaux de H.

La forme échelonnée réduite de la matrice A donne simultanément une base pour l'espace H et une base pour $H^\perp$. Géométriquement, la base pour H c'est les directions permises ; la base pour $H^\perp$ c'est les directions défendues.

H possède trois vecteurs directeurs et deux vecteurs normales. Les vecteurs ne sont pas unique (il y a une infinité de bases pour $\mathrm{col}(A^T)$ et $\ker(A)$), mais le nombre est ! Comme contraste, un plan dans $\mathbb{R}^3$ possède deux vecteurs directeurs et un vecteur normal ; une droite dans $\mathbb{R}^2$ possède un vecteur directeur et un vecteur normal.

[1] ou si on préfère les trois rangées finales

4. Espaces Orthogonaux et Matrices II

Nos observations de la première section permettent aussi de conclure que

> **Théorème 6.** $\ker(B^T) = \mathrm{col}(B)^{\perp}$

Il s'agit simplement de $B = A^T$.

Exemple: Trouver les conditions pour $\mathrm{col}(B)$, où $B = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

$$\left[\begin{array}{cc|c} -2 & -3 & x \\ 1 & 0 & y \\ 0 & 1 & z \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{cc|c} 1 & 0 & y \\ 0 & 1 & z \\ 0 & 0 & x+2y+3z \end{array}\right]$$

Un vecteur $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ est dans $\mathrm{col}(A)$ exactement lorsque $x + 2y + 3z = 0$. Plutôt :

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathrm{col}(B) \iff \begin{bmatrix} x \\ y \\ z \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 0$$

Alors on a que $\ker(B^T) = \left\{ t \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \mid t \in \mathbb{R} \right\}$. Une base pour $\ker(B^T)$ est $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$. On voit de plus que les colonnes de B forment une base pour l'espace P de la première section, car il y a un pivot dans chaque colonne.

> **Algorithme 7.** *Soit $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ des vecteurs dans $\mathbb{R}^n$, et U l'espace qu'ils engendrent. Afin de trouver une base pour $U^{\perp}$:*
>
> *(1) On met les vecteurs en colonne dans une matrice augmentée B.*
>
> *(2) On réduit la matrice à la forme échelonnée afin de trouver les conditions pour $\mathrm{col}(B)$.*
>
> *(3) Pour chaque condition, on prend les coefficients pour former un vecteur : c'est une base.*

Comme auparavant, en trouvant la base pour $\ker(B^T)$, on trouve une base pour $\mathrm{col}(B)$ aussi.

5. Un Exemple II

Exemple: Soit H l'espace engendré par $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 5 \\ 1 \\ 2 \end{bmatrix} \right\}$

Trouver une base pour H, et une base pour $H^{\perp}$.

On met les vecteurs en colonne, et on réduit :

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & x_1 \\ 0 & 1 & 1 & x_2 \\ 2 & 3 & 5 & x_3 \\ 0 & 1 & 1 & x_4 \\ 1 & 0 & 2 & x_5 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 1 & x_1 \\ 0 & 1 & 1 & x_2 \\ 0 & 0 & 1 & -x_1 + x_5 \\ 0 & 0 & 0 & -x_2 + x_4 \\ 0 & 0 & 0 & -2x_1 - 3x_2 + x_3 \end{array}\right]$$

Il y a un pivot dans chaque colonne. Donc les trois colonnes originaux[2] forment une base pour $\mathrm{col}(B)$. Donc les trois vecteurs originaux forment une base pour H. Si on avait eu deux pivots, on aurait retenu seulement deux vecteurs pour former la base (et alors H serait un plan en $\mathbb{R}^5$)

Les conditions sont :

$$0 = -1x_1 + 0x_2 + 0x_3 + 1x_4 + 0x_5 \qquad 0 = -2x_1 - 3x_2 + 1x_3 + 0x_4 + 0x_5$$

On prend les coefficients pour donner une base pour $\ker(B^T)$:

$$\left\{ \begin{bmatrix} -2 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \right\}$$

Ce sont *deux* vecteurs normaux de H.

La forme échelonnée de la matrice B donne simultanément une base pour l'espace H et une base pour $H^{\perp}$. Géométriquement, la base pour H c'est les directions permises ; la base pour $H^{\perp}$ c'est les directions défendues. On voit aussi que les direction défendues sont exactement les conditions pour H : on pourrait dire que l'équation de H est "$0 = -1x_1 + 0x_2 + 0x_3 + 1x_4 + 0x_5,\ 0 = -2x_1 - 3x_2 + 1x_3 + 0x_4 + 0x_5$"

6. Résumé

C'est beaucoup de concepts, toutes mises ensemble. Mais le résultat est directe, et ne requiert rien de plus que la réduction de Gauss-Jordan

[2] et non les trois colonnes finales

La forme échelonnée réduite d'une matrice A donne simultanément une base pour l'espace rangée $\text{col}(A^T)$ et son complément orthogonal $\ker(A)$. Ce sont les directions permises et directions défendues de "l'hyper-plan" engendré par les rangées de A.

La forme échelonnée d'une matrice B donne simultanément une base pour l'espace colonne $\text{col}(B)$ et son complément orthogonal $\ker(B^T)$. Ce sont les directions permises et directions défendues de "l'hyper-plan" engendré par les colonnes de B.

On a donc deux méthodes pour trouver une base de $U^\perp$, sachant des vecteurs qui engendrent U : soit mettre les vecteurs en rangée dans une matrice A ou en colonne dans une matrice B.

De plus, on connît que les conditions d'un espace donnent une base pour son complément orthogonal. En langage plus simple : les coefficients de l'équation d'un plan sont exactement un vecteur normal — valide dans n'importe quelle dimension.

Exemple: Soit la matrice M. Donner tout.

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 2 & x_1 \\ 1 & 1 & 3 & 4 & x_2 \\ 1 & -1 & -1 & 0 & x_3 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|l} 1 & 0 & 1 & 2 & x_1 \\ 0 & 1 & 2 & 2 & -x_1 + x_2 \\ 0 & 0 & 0 & 0 & -2x_1 + x_2 + x_3 \end{array}\right]$$

Une base pour $\text{col}(M)$: on retient les colonnes originales avec pivot, donc $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$.

Une base pour $\text{col}(M^T)$: les rangées avec pivot, donc $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 3 \\ 4 \end{bmatrix} \right\}$ ou $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 2 \end{bmatrix} \right\}$

Une base pour $\ker(M)$: on solutionne avec $a = b = c = 0$, donc $\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Une base pour $\ker(M^T)$: les conditions sur l'espace colonne, donc $\left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Le rang est 2. Les dimensions de ces quatre espaces sont 2, 2, 2, 1. On vérifie que ces quatre chiffres sont $\text{rang}(M)$, $\text{rang}(M)$, $n - \text{rang}(M)$, $m - \text{rang}(M)$.

Il n'y a pas de pivot dans chaque rangée. Les rangées sont dépendantes. Les colonnes n'engendrent pas $\mathbb{R}^3$. Donc $\text{col}(M) = \{\mathbb{R}^3 \mid 0 = -2x_1 + x_2 + x_3\}$.

Il n'y a pas de pivot dans chaque colonne. Les colonnes sont dépendantes. Les rangées n'engendrent pas $\mathbb{R}^4$. La base pour $\ker(M)$ donne les conditions sur l'espace engendré. Donc $\text{col}(M^T) = \{\mathbb{R}^4 \mid 0 = -y_1 - 2y_2 + y_3,\ 0 = -2y_1 - 2y_2 + y_4\}$.

7. Un Défi

C'est optionnelle, mais pour ceux qui veulent comprendre plus de la théorie des espaces orthogonaux c'est peut-être intéressant.

(1) C'est parfois lourd de faire des réductions avec des inconnues à la droite. Est-ce possible de trouver une autre méthode d'accomplir la même chose ?

(2) Si on cherche l'inverse d'une matrice en réduisant $[A|I]$, on trouve parfois qu'il y a moins de n pivots, donc pas d'inverse. Qu'est-ce que c'est la vraie signifiance des rangées nulles ?

(3) Si on retient seulement les rangées de A avec pivot, il existe une inverse à droite. Ou est-elle ? Il y a une infinité de tels inverses. Ou sont-elles ?

Une indice :

$$\left[\begin{array}{cccc|ccc} 1 & 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 3 & 4 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 0 & 1 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|ccc} 1 & 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -2 & 1 & 1 \end{array}\right]$$

1. Espaces Orthogonaux et Vecteurs

Rappelons le plan $P = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x + 2y + 3z = 0 \right\}$; on connaît la base $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$.
C'est donc l'espace colonne de la matrice $A = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

On pourrait demander si, par exemple, le point $\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$ est sur le plan P. On voudrait alors savoir si le système suivant est consistant :

$$\left[\begin{array}{cc|c} -2 & -3 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right]$$

Exercice: Montrer que ce n'est pas consistant. De plus vérifier que la condition $x + 2y + 3z = 0$ n'est pas satisfaite.

Savoir que ce n'est pas consistant ne finit pas l'affaire. On pourrait demander de *décomposer* le vecteur selon P. Précisément, trouver un vecteur $\mathbf{b}_1 \in P$ et $\mathbf{b}_2 \in P^\perp$ tel que $\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2$. On comprend ceci de façon intuitive : le vecteur $\mathbf{b}_1$ est la partie de $\mathbf{b}$ qui est dans P, le vecteur $\mathbf{b}_2$ est la partie de $\mathbf{b}$ qui est orthogonale à P.

2. Projections

On voudrait que $A\mathbf{x} = \mathbf{b}$ possède une solution ; malheureusement, la réponse est non (l'exercice !). Donc on cherche une décomposition $\mathbf{b}_1 + \mathbf{b}_2$, comme ci-haut.

Imaginons qu'on connaît une solution *approximative* $\mathbf{x}$, tel que $A\mathbf{x}$ est aussi proche à $\mathbf{b}$ que possible. Géométriquement, ceci veut dire que le vecteur dirigé de $A\mathbf{x}$ à $\mathbf{b}$ est orthogonal au plan P. Mais orthogonal au plan P veut dire dans $P^\perp$. Donc on voudrait que $A\mathbf{x} - \mathbf{b}$ soit dans $\ker(A^T)$. Un dessin est essentiel pour bien comprendre.

On ne connaît encore pas $\mathbf{x}$! Mais on peut la déterminer en exigeant que $A\mathbf{x} - \mathbf{b} \in \ker(A^T)$. On exige donc $A^T (A\mathbf{x} - \mathbf{b}) = \mathbf{0}$. Ceci est équivalent à

$$A^T A\mathbf{x} = A^T\mathbf{b}$$

En solutionnant pour $\mathbf{x}$, on trouve que $A\mathbf{x}$ est le point sur le plan qui est le plus proche de $\mathbf{b}$ (c'est comme ça qu'on a définit $\mathbf{x}$).

C'est le moment de mettre des chiffres afin de voir ce qui ce passe.

On a $A = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ et $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Donc

$$A^T A = \begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 6 & 10 \end{bmatrix} \qquad A^T\mathbf{b} = \begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

On cherche alors à solutionner

$$\left[\begin{array}{cc|c} 5 & 6 & -1 \\ 6 & 10 & -2 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & \frac{1}{7} \\ 0 & 1 & -\frac{2}{7} \end{array}\right]$$

La solution est $x_1 = \frac{1}{7}$ et $x_2 = -\frac{2}{7}$. Donc la projection est

$$A\mathbf{x} = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{7} \\ -\frac{2}{7} \end{bmatrix} = \begin{bmatrix} \frac{4}{7} \\ \frac{1}{7} \\ -\frac{2}{7} \end{bmatrix}$$

Noter que $\mathbf{x}$ n'est PAS la projection : ce n'est même pas un vecteur de la bonne taille ! On cherche un point dans l'espace $\mathbb{R}^3$. Le vecteur $\mathbf{x}$ donne la combinaison linéaire de la base pour P.

De plus on a

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2 = \begin{bmatrix} \frac{4}{7} \\ \frac{1}{7} \\ -\frac{2}{7} \end{bmatrix} + \left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{4}{7} \\ \frac{1}{7} \\ -\frac{2}{7} \end{bmatrix} \right) = \begin{bmatrix} \frac{4}{7} \\ \frac{1}{7} \\ -\frac{2}{7} \end{bmatrix} + \begin{bmatrix} \frac{3}{7} \\ \frac{6}{7} \\ \frac{9}{7} \end{bmatrix}$$

avec $\mathbf{b}_1 \in U$ et $\mathbf{b}_2 \in U^\perp$.

On peut montrer que $\operatorname{rang}(A^T A) = \operatorname{rang}(A^T A | A^T\mathbf{b})$. La conséquence c'est que ce système serait toujours consistant (on se rappelle d'un des Quatre Théorèmes). Ce ne serait pas nécessairement unique. Par contre, pour n'importe quelle solution $\mathbf{x}$, on obtiendrait toujours la même projection $A^T\mathbf{x}$.

Algorithme 1. *Afin de trouver la projection de* $\mathbf{b}$ *sur un espace* U *:*

(1) Écrire une matrice A tel que les colonnes de A engendrent U. Ce n'est pas nécessaire que les colonnes forment une base.

(2) Solutionner le système $A^T A\mathbf{x} = A^T\mathbf{b}$ *pour* $\mathbf{x}$.

(3) La projection est $\operatorname{proj}_U(\mathbf{b}) = A\mathbf{x}$.

(4) On a $\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2$ *avec* $\mathbf{b}_1 = \operatorname{proj}_U(\mathbf{b}) \in U$ *et* $\mathbf{b}_2 = \mathbf{b} - \operatorname{proj}_U(\mathbf{b}) \in U^\perp$.

3. Matrice de Projection : un Aperçu

Dans l'exemple ci-haut les colonnes de A forment une base pour P. C'est un cas spécial ! Donc la matrice A possède un pivot dans chaque colonne ; autrement dit, $\ker(A) = \{\mathbf{0}\}$. Mais alors on connaît que $A^T A$ est inversible (c'était un devoir). Donc le système $A^T A\mathbf{x} = A^T\mathbf{b}$ est consistant et possède une solution unique. De plus, on la connaît :

$$\mathbf{x} = \left(A^T A\right)^{-1} A^T\mathbf{b}$$

Afin de calculer la projection, on veut $A\mathbf{x}$, donc

$$A\mathbf{x} = A\left(A^T A\right)^{-1} A^T\mathbf{b}$$

Soit $M = A\left(A^T A\right)^{-1} A^T$; la projection de $\mathbf{b}$ sur $\operatorname{col}(A)$ est exactement $M\mathbf{b}$.

Donc, pour notre plan P on aura :

$$A^T A = \begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 6 & 10 \end{bmatrix}$$

$$\left(A^T A\right)^{-1} = \cdots = \begin{bmatrix} \frac{5}{7} & -\frac{3}{7} \\ -\frac{3}{7} & \frac{5}{14} \end{bmatrix}$$

$$M = A\left(A^T\right)^{-1} A^T = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{5}{7} & -\frac{3}{7} \\ -\frac{3}{7} & \frac{5}{14} \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{13}{14} & -\frac{1}{7} & -\frac{3}{14} \\ -\frac{1}{7} & \frac{5}{7} & -\frac{3}{7} \\ -\frac{3}{14} & \frac{3}{7} & \frac{5}{14} \end{bmatrix}$$

Maintenant on calcul

$$M\mathbf{b} = \begin{bmatrix} \frac{13}{14} & -\frac{1}{7} & -\frac{3}{14} \\ -\frac{1}{7} & \frac{5}{7} & -\frac{3}{7} \\ -\frac{3}{14} & \frac{3}{7} & \frac{5}{14} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{4}{7} \\ \frac{1}{7} \\ -\frac{2}{7} \end{bmatrix}$$

Algorithme 2. *Afin de trouver la projection de* $\mathbf{b}$ *sur un espace* U *:*

(1) Écrire une matrice A tel que les colonnes de A forment une base pour U.
(2) Calculer la matrice $M = A\left(A^T A\right)^{-1} A^T$.
(3) La projection est $\operatorname{proj}_U(\mathbf{b}) = M\mathbf{b}$.
(4) On a $\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2$ avec $\mathbf{b}_1 = \operatorname{proj}_U(\mathbf{b}) \in U$ et $\mathbf{b}_2 = \mathbf{b} - \operatorname{proj}_U(\mathbf{b}) \in U^\perp$.

Cette méthode est plus longue que la première. Par contre, si on a beaucoup de $\mathbf{b}$, on économise au long.

NB: Attention ! Ce n'est pas vrai que $A\left(A^T A\right)^{-1} A^T = AA^{-1}\left(A^T\right)^{-1} A^T = II = I$. La matrice A n'est peut-être pas carré. Par contre si $\ker(A) = \{\mathbf{0}\}$ alors la matrice (carrée !) $A^T A$ est inversible.

4. Un Cas Spécial

Posons $D = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Alors col$(D)$ est une droite dans $\mathbb{R}^2$. Soit le point $\mathbf{b} = \begin{bmatrix} 3 & 4 \end{bmatrix}$ dans $\mathbb{R}^2$. On voit que $\mathbf{b}$ n'est pas sur la droite D. Trouvons alors la projection de $\mathbf{b}$ sur D.

On résoud premièrement $D^T D\mathbf{x} = D^T\mathbf{b}$:

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Note que ici $\mathbf{x}$ est un vecteur de taille 1×1, donc un chiffre. Aussi, la multiplication matricielle de D^T et D correspond au produit scalaire de D (la colonone!) avec lui-même.

$$\mathbf{x} = x = \frac{\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 4 \end{bmatrix}}{\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \end{bmatrix}}$$

Donc la projection est

$$D\mathbf{x} = xD = \frac{\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 4 \end{bmatrix}}{\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \end{bmatrix}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

C'est exactement la formule pour la projection d'un vecteur sur une autre qu'on a vu au chapitre 2.

5. Approximation de Moindres Carrées

C'est une section optionnelle : on découvre que les approximations de moindres carrées reposent sur l'algèbre linéaire.

On a plusieurs points sur un plan et on cherche la droite qui est la meilleure approximation à ces points.

Faisons un exemple. Voici quatre points sur un plan. Faites un graphique : ils ne sont pas alignés ; proches mais pas parfait.

$$(2,1) \quad (7,3) \quad (5,2) \quad 10,4$$

On cherche une droite de la forme $y = mx + b$ tel que les points sont tous proches à la droite. Donc on voudrait que

$$\begin{aligned} b + m(2) &= (1) \\ b + m(7) &= (3) \\ b + m(5) &= (2) \\ b + m(10) &= (4) \end{aligned}$$

On reconnaît un système linéaire à quatre équations et deux variables. Posons A pour la matrice des coefficients et $\mathbf{y}$ pour les constantes (car les constantes à droite sont les valeurs-y des points). Les variables sont b et m :

$$\begin{bmatrix} 1 & 2 \\ 1 & 7 \\ 1 & 5 \\ 1 & 10 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} \qquad \text{donnant} \qquad \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 7 & 3 \\ 1 & 5 & 2 \\ 1 & 10 & 4 \end{array}\right]$$

Il n'y a pas de solution : aucune droite ne passe exactement par ces quatre points. Mais on peut trouver une approximation en solutionnant $A^T A \begin{bmatrix} b \\ m \end{bmatrix} = A^T \mathbf{y}$. Donc

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 7 & 5 & 10 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 7 \\ 1 & 5 \\ 1 & 10 \end{bmatrix} = \begin{bmatrix} 4 & 24 \\ 24 & 178 \end{bmatrix}$$

$$A^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 7 & 5 & 10 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 73 \end{bmatrix}$$

On solutionne

$$\left[\begin{array}{cc|c} 4 & 24 & 10 \\ 24 & 178 & 73 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & \frac{7}{34} \\ 0 & 1 & \frac{13}{34} \end{array}\right]$$

qui donne $b = \frac{7}{34}$ et $m = \frac{13}{34}$.

La droite optimale est alors $y = \frac{13}{34}x + \frac{7}{34}$.

On peut adapter cetter méthode pour trouver la meilleure approximation en forme de quadratique, cubique, ou même d'autres formes d'équations.

MAT 1741 — **20. Bases Orthogonaux et Orthonormaux**

Mike Newman — Université d'Ottawa — automne 2008

1. Bases

C'est souvent utile d'avoir une base ou touts les vecteurs sont perpendiculaires. Aussi, que chaque vecteur dans la base est de norme un.

> **Définition 1.** *Une* BASE ORTHOGONALE $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est une base tel que* $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ *lorsque* $i \neq j$.
>
> *Une* BASE ORTHONORMALE $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *est une base tel que* $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ *lorsque* $i \neq j$ *et que* $\|\mathbf{x}_i\| = 1$.

Exemple: $\left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\}$ est une base orthonormale pour $\mathbb{R}^3$.

Les produits scalaires sont tous zéro :

$$\begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}0\\1\\0\end{bmatrix} = 0+0+0 = 0$$
$$\begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = 0+0+0 = 0$$
$$\begin{bmatrix}0\\1\\0\end{bmatrix} \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = 0+0+0 = 0$$

et de plus les normes sont tous un :

$$\left\| \begin{bmatrix}1\\0\\0\end{bmatrix} \right\| = \begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = \sqrt{(1)^2+(0)^2+(0)^2} = 1$$
$$\left\| \begin{bmatrix}0\\1\\0\end{bmatrix} \right\| = \begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}0\\1\\0\end{bmatrix} = \sqrt{(0)^2+(1)^2+(0)^2} = 1$$
$$\left\| \begin{bmatrix}0\\0\\1\end{bmatrix} \right\| = \begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = \sqrt{(0)^2+(0)^2+(1)^2} = 1$$

Exemple: $\left\{ \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix} \right\}$ est une base orthogonale pour $\mathbb{R}^2$, mais pas orthonormale.

Il n'y a qu'un produit scalaire à vérifier :

$$\begin{bmatrix}1\\1\end{bmatrix} \cdot \begin{bmatrix}1\\-1\end{bmatrix} = 1-1 = 0$$

C'est zéro, donc c'est une base orthogonale. Par contre les normes sont

$$\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\| = \sqrt{(1)^2 + (1)^2} = \sqrt{2}$$
$$\left\| \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\| = \sqrt{(1)^2 + (-1)^2} = \sqrt{2}$$

Exemple: Montrer que $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$ n'est pas une base orthogonale pour $\mathbb{R}^3$.

On peut calculer que chaque produit scalaire est bien zéro. Mais ce n'est pas une base : le vecteur zéro ne peut jamais faire partie d'une base, car l'ensemble serait dépendant. On se rappelle (début du cours !) que le seul vecteur qui a norme 0 est le vecteur zéro ; autrement dit, $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$. Donc

Proposition 2.

Soit $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *une base orthogonale. Alors* $\|\mathbf{x}_i\| > 0$ *pour chaque* $\mathbf{x}_i$.

L'intuition géométrique suggère que les vecteurs orthogonaux ne pourront pas être dépendants.

Théorème 3.

Soit $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *des vecteurs orthogonaux. Alors ils sont indépendants.*

Pour prouver ceci, supposons qu'on a

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n = \mathbf{0}$$

On voudrait montrer que la seule solution pour les a_i est la solution triviale. On prend le produit scalaire avec $\mathbf{x}_1$ pour donner

$$\begin{aligned} 0 = \mathbf{x}_1 \cdot (\mathbf{0}) &= \mathbf{x}_1 \cdot (a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n) \\ &= a_1(\mathbf{x}_1 \cdot \mathbf{x}_1) + a_2(\mathbf{x}_1 \cdot \mathbf{x}_2) + \cdots + a_n(\mathbf{x}_1 \cdot \mathbf{x}_n) \\ &= a_1(\mathbf{x}_1 \cdot \mathbf{x}_1) + a_2(0) + \cdots + a_n(0) \\ &= a_1(\mathbf{x}_1 \cdot \mathbf{x}_1) \\ &= a_1\|\mathbf{x}_1\|^2 \end{aligned}$$

Puisque $\|\mathbf{x}_1\| \neq 0$, on conclut que $a_1 = 0$. De la même manière, chaque $a_i = 0$, et nécessairement l'ensemble est indépendant. $\square$

C'est très utile. Par exemple, on connaît automatiquement qu'un ensemble de n vecteurs orthogonaux dans $\mathbb{R}^n$ est une base pour $\mathbb{R}^n$. En particulier, la deuxième exemple ci-haut est une base pour $\mathbb{R}^2$ (c'est évident pour d'autres raisons).

2. Projection et Coordonnées

On a déja vu comment trouver la projection d'un vecteur sur U et $U^\perp$, pour un sous-espace U. On peut maintenant faire une décomposition d'un vecteur par rapport à une base prthogonale.

Soit $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ une base orthogonale pour $\mathbb{R}^n$, et $\mathbf{y}$ n'importe quel vecteur dans $\mathbb{R}^n$. On connaaît que $\mathbf{y}$ sécrit de manière unique comme combinaison linéaire de cette base. Puisque la base est orthogonale, c'est facile à calculer : on calcul les projections sur chaque $\mathbf{x}_i$.

Théorème 4. *Soit $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ une base orthogonale pour $\mathbb{R}^n$, et $\mathbf{y} \in \mathbb{R}^n$. Alors*

$$\mathbf{y} = \frac{\mathbf{x}_1 \cdot \mathbf{y}}{\|\mathbf{x}_1\|^2}\mathbf{x}_1 + \frac{\mathbf{x}_2 \cdot \mathbf{y}}{\|\mathbf{x}_2\|^2}\mathbf{x}_2 + \cdots + \frac{\mathbf{x}_n \cdot \mathbf{y}}{\|\mathbf{x}_n\|^2}\mathbf{x}_n$$

Les valeurs

$$\frac{\mathbf{x}_1 \cdot \mathbf{y}}{\|\mathbf{x}_1\|^2}, \quad \frac{\mathbf{x}_2 \cdot \mathbf{y}}{\|\mathbf{x}_2\|^2}, \quad \ldots, \quad \frac{\mathbf{x}_n \cdot \mathbf{y}}{\|\mathbf{x}_n\|^2}$$

sont les COORDONNÉES *de $\mathbf{y}$ par rapport à cette base.*

Note que si c'est une base orthonormale, alors les dénominateurs sont tous égal à 1.

Exemple: Trouver les coordonnées de $\begin{bmatrix}3\\4\end{bmatrix}$ par rapport à la base $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right\}$.

Puisque c'est une base orthogonale, on peut utiliser la formule.

$$\frac{\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] \cdot \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right]}{\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] \cdot \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]} = \frac{7}{2}$$

$$\frac{\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right] \cdot \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right]}{\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right] \cdot \left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]} = \frac{-1}{2}$$

Alors on a que

$$\begin{bmatrix}3\\4\end{bmatrix} = \frac{7}{2}\begin{bmatrix}1\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}\frac{7}{2}\\\frac{7}{2}\end{bmatrix} + \begin{bmatrix}-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$$

3. Bases Orthogonales et Projection

Supposons qu'on a un sous-espace U dans $\mathbb{R}^n$. C'est utile de connaître une base pour U, mais pourquoi est-ce utile de connaître une base *orthogonale* ?

Une raison c'est pour calculer des projections. Afin de faire une projection sur U, on a besoin d'un ensemble de vecteurs qui engendrent U : voir le chapitre précédent. Mais supposons que les colonnes de A forment une base orthogonale pour U. C'est certainement un ensemble engendrant pour U, donc on peut calculer la projection comme auparavant.

Précisément : on suppose que $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ est une base orthogonale pour U, qui est sous-espace de $\mathbb{R}^m$, et que $\mathbf{b}$ est un vecteur dans $\mathbb{R}^m$. Comme mini-exercice, montrer que $\left(A^T A\right)_{ij}$ est exactement $\mathbf{u}_i \cdot \mathbf{u}_j$ (revoir la définition du produit matricielle $A^T A$). C'est parce qu'on prend le produit scalaire de la i-ième rangée de A^T avec la j-ième colonne de A. Donc $A^T A$ est diagonale, et de plus les éléments sur le diagonal sont exactement $\mathbf{u}_i \cdot \mathbf{u}_i = \|\mathbf{u}_i\|^2$. Donc le système linéaire $A^T A\mathbf{x} = A^T\mathbf{b}$ est exactement le système

$$\begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{u}_1 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{u}_2 \cdot \mathbf{u}_2 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{u}_3 \cdot \mathbf{u}_3 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \mathbf{u}_n \cdot \mathbf{u}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{b} \\ \mathbf{u}_2 \cdot \mathbf{b} \\ \mathbf{u}_3 \cdot \mathbf{b} \\ \vdots \\ \mathbf{u}_n \cdot \mathbf{b} \end{bmatrix}$$

C'est un système linéaire où la matrice des coefficients est diagonale. En forme de système d'équations on a :

$$\begin{aligned} (\mathbf{u}_1 \cdot \mathbf{u}_1)\, x_1 &= (\mathbf{u}_1 \cdot \mathbf{b}) \\ (\mathbf{u}_2 \cdot \mathbf{u}_2)\, x_2 &= (\mathbf{u}_2 \cdot \mathbf{b}) \\ &\vdots \\ (\mathbf{u}_n \cdot \mathbf{u}_n)\, x_n &= (\mathbf{u}_n \cdot \mathbf{b}) \end{aligned}$$

Les valeurs entre paranthèses sont des chiffres. On résoud le système en isolant chaque équation pour x_i, donnant la solution :

$$x_1 = \frac{\mathbf{u}_1 \cdot \mathbf{b}}{\mathbf{u}_1 \cdot \mathbf{u}_1}, \quad x_2 = \frac{\mathbf{u}_2 \cdot \mathbf{b}}{\mathbf{u}_2 \cdot \mathbf{u}_2}, \quad \cdots \quad x_n = \frac{\mathbf{u}_n \cdot \mathbf{b}}{\mathbf{u}_n \cdot \mathbf{u}_n}$$

On vient de découvrir une formule pour calculer la projection sur une espace U, sachant une base orthonormale pour U.

Théorème 5. *Soit* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ *une base orthogonale pour un sous-espace* U *de* $\mathbb{R}^n$*, et* $\mathbf{y} \in \mathbb{R}^n$*.*

Alors

$$\mathrm{proj}_U(\mathbf{y}) = \frac{\mathbf{x}_1 \cdot \mathbf{y}}{\|\mathbf{x}_1\|^2}\mathbf{x}_1 + \frac{\mathbf{x}_2 \cdot \mathbf{y}}{\|\mathbf{x}_2\|^2}\mathbf{x}_2 + \cdots + \frac{\mathbf{x}_m \cdot \mathbf{y}}{\|\mathbf{x}_m\|^2}\mathbf{x}_m$$

Note que c'est presque la même chose que la formule pour trouver les coordonnées d'un vecteur par rapport à une base orthogonale. En fait la Théorème 4 est un cas spécial, car $\mathbb{R}^m$ est toujours un sous-espace de $\mathbb{R}^m$.

Exercice: Supposons que dans les Théorèmes 4 et 5 on a une base *orthonormale*. Comment est-ce que la formule se simplifie ? De plus, dans la discussion ci-haut, qu'est-ce que c'est la matrice A^TA dans le cas ou les colonnes de A forment une base orthonormale pour U ?

4. Gram-Schmidt

Il reste la question de comment trouver des bases orthogonaux ? L'idée clé c'est que si deux vecteurs ne sont pas orthogonaux, on peut enlever la projection de l'un sur l'autre pour donner deux vecteurs orthogonaux.

Algorithme 6. *Soit* $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ *une base pour un espace* U*.*

Afin de trouver une base orthogonale pour U *on suit répète le suivant, avec* $\mathbf{v}_i = \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ *:*

- *Pour chaque autre vecteur* $\mathbf{v}_j$ *avec* $j > i$*, on enlève la projection de* $\mathbf{v}_j$ *sur* $\mathbf{v}_i$*. C'est-à-dire* $\mathbf{v}_j \mapsto \frac{\mathbf{v}_j \cdot \mathbf{v}_j}{\|\mathbf{v}_i\|^2}$*.*

On a maintenant une base orthogonale. Afin d'obtenir une base orthonormale

- *Diviser chaque vecteur par sa norme.*

Un exemple est essentiel pour comprendre !

Exemple: Utiliser la méthode de Gram-Schmidt pour transformer la base $\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}$ pour $\mathbb{R}^3$ en base orthonormale pour $\mathbb{R}^3$.

On commence avec le premier vecteur, et on ajuste les deux autres afin qu'ils deviennent orthogonaux au premier.

On ajuste le deuxième vecteur relatif au premier. La projection du deuxième sur le premier est :

$$\frac{\begin{bmatrix}0\\1\\1\end{bmatrix}\cdot\begin{bmatrix}1\\1\\0\end{bmatrix}}{\begin{bmatrix}0\\1\\1\end{bmatrix}\cdot\begin{bmatrix}0\\1\\1\end{bmatrix}}\begin{bmatrix}0\\1\\1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$$

On soustrait pour donner le nouveau deuxième vecteur :

$$\begin{bmatrix}1\\1\\0\end{bmatrix}-\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}=\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$$

On ajuste le troisième vecteur relatif au premier. La projection du troisième sur le premier est :

$$\frac{\begin{bmatrix}0\\1\\1\end{bmatrix}\cdot\begin{bmatrix}0\\1\\2\end{bmatrix}}{\begin{bmatrix}0\\1\\1\end{bmatrix}\cdot\begin{bmatrix}0\\1\\1\end{bmatrix}}\begin{bmatrix}0\\1\\1\end{bmatrix}=\frac{3}{2}\begin{bmatrix}0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\\frac{3}{2}\\\frac{3}{2}\end{bmatrix}$$

On soustrait pour donner le nouveau troisième vecteur :

$$\begin{bmatrix}0\\1\\2\end{bmatrix}-\begin{bmatrix}0\\\frac{3}{2}\\\frac{3}{2}\end{bmatrix}=\begin{bmatrix}0\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$$

On a maintenant fini d'ajuster relatif au premier vecteur. À ce moment la base est

$$\left\{\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix},\begin{bmatrix}0\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}\right\}$$

Le premier vecteur est orthogonal aux autres. On peut ignorer le premier vecteur, et maintenant on peut passer au deuxième vecteur.

On ajuste le troisième vecteur relatif au deuxième. Note que on prend les « nouveau » vecteurs. On ne veut absoluement *pas* reprendre les vecteurs au début. C'est similaire à la méthode Gauss-Jordan : on s'en sert toujours des nouvelles rangées, laissant de côté les rangées originales. En tout cas, la projection du troisième sur le deuxième est :

$$\frac{\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}\cdot\begin{bmatrix}0\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}}{\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}\cdot\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}}\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}=\frac{-\frac{1}{2}}{\frac{3}{2}}\begin{bmatrix}1\\\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}=\begin{bmatrix}-\frac{1}{3}\\-\frac{1}{6}\\\frac{1}{6}\end{bmatrix}$$

On soustrait pour donner le nouveau nouveau (!) troisième vecteur :

$$\begin{bmatrix}0\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}-\begin{bmatrix}-\frac{1}{3}\\-\frac{1}{6}\\\frac{1}{6}\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\-\frac{1}{3}\\\frac{1}{3}\end{bmatrix}$$

On a maintenant fini d'ajuster relatif au deuxième vecteur. À ce moment on a une base orthogonale : il n'y a plus rien à ajuster.

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \right\}$$

Afin d'obtenir une base orthonormale, on calcul les normes des vecteurs :

$$\left\| \left[\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix} \right] \right\| = \sqrt{(0)^2 + (1)^2 + (1)^2} = \sqrt{2}$$

$$\left\| \left[\begin{smallmatrix} 1 \\ \frac{1}{2} \\ -\frac{1}{2} \end{smallmatrix} \right] \right\| = \sqrt{(1)^2 + (\frac{1}{2})^2 + (-\frac{1}{2})^2} = \sqrt{\frac{3}{2}}$$

$$\left\| \left[\begin{smallmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \end{smallmatrix} \right] \right\| = \sqrt{(\frac{1}{3})^2 + (-\frac{1}{3})^2 + (\frac{1}{3})^2} = \sqrt{\frac{1}{3}}$$

On divise chaque vecteur par sa norme afin d'obtenir une base orthonormale :

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{2}} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad \begin{bmatrix} 1 \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \sqrt{\frac{2}{3}} = \begin{bmatrix} \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \sqrt{3} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

La base orthonormale est

$$\left\{ \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \right\}$$

Exercice: vérifier que les produits scalaires de différents vecteurs dans cette base sont tous zéro, et que les normes sont tous un.

Exemple: La suivante est une base pour un espace U. La trasformer en base orthonormale pour U, suivant la méthode de Gram-Schmidt. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} \right\}$

C'est une base pour un sous-espace U de $\mathbb{R}^4$. On ne connaît pas U (sauf que c'est engendré par cette base). On pourraît imaginer qu'on a trouvé cette base à l'aide de la méthode Gauss-Jordan ; par exemple, en cherchant une base pour en espace engendré par un ensemble de vecteurs. Maintenant on cherche une base *orthonormale* pour ce même espace.

On commence en ajustant pour que touts les autres vecteurs deviennent orthogonaux au premier.

La projection de le deuxième sur le premier :

$$\frac{\begin{bmatrix}\frac{1}{2}\\1\\0\end{bmatrix}\cdot\begin{bmatrix}0\\1\\1\\1\end{bmatrix}}{\begin{bmatrix}\frac{1}{2}\\1\\0\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{2}\\1\\0\end{bmatrix}}\begin{bmatrix}1\\2\\1\\0\end{bmatrix}=\frac{3}{6}\begin{bmatrix}1\\2\\1\\0\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\\1\\\frac{1}{2}\\0\end{bmatrix}$$

On soustrait pour obtenir le nouveau deuxième vecteur :

$$\begin{bmatrix}0\\1\\1\\1\end{bmatrix}-\begin{bmatrix}\frac{1}{2}\\1\\\frac{1}{2}\\0\end{bmatrix}=\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}$$

La projection de le troisième sur le premier :

$$\frac{\begin{bmatrix}\frac{1}{2}\\1\\0\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{2}\\1\\1\end{bmatrix}}{\begin{bmatrix}\frac{1}{2}\\1\\0\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{2}\\1\\0\end{bmatrix}}\begin{bmatrix}1\\2\\1\\0\end{bmatrix}=\frac{6}{6}\begin{bmatrix}1\\2\\1\\0\end{bmatrix}=\begin{bmatrix}1\\2\\1\\0\end{bmatrix}$$

On soustrait pour obtenir le nouveau troisième vecteur :

$$\begin{bmatrix}1\\2\\1\\1\end{bmatrix}-\begin{bmatrix}1\\2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

On a fini d'ajuster relatif au premier ; on passe au deuxième. Il n'y qu'un ajustement à faire : il faut que le troisième soit orthogonal au deuxième.

La projection de le troisième sur le deuxième (note qu'on utilise les nouveau vecteurs, pas les originaux) :

$$\frac{\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}\cdot\begin{bmatrix}0\\0\\0\\1\end{bmatrix}}{\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}\cdot\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}}\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}=\frac{1}{\frac{3}{2}}\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}=\begin{bmatrix}-\frac{1}{3}\\0\\\frac{1}{3}\\\frac{2}{3}\end{bmatrix}$$

On soustrait pour obtenir le nouveau troisième vecteur :

$$\begin{bmatrix}0\\0\\0\\1\end{bmatrix}-\begin{bmatrix}-\frac{1}{3}\\0\\\frac{1}{3}\\\frac{2}{3}\end{bmatrix}=\begin{bmatrix}\frac{1}{3}\\0\\-\frac{1}{3}\\\frac{1}{3}\end{bmatrix}$$

On a maintenant une base orthogonale pour U : $\left\{\begin{bmatrix}1\\2\\1\\0\end{bmatrix},\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix},\begin{bmatrix}\frac{1}{3}\\0\\-\frac{1}{3}\\\frac{1}{3}\end{bmatrix}\right\}$

On calcul les normes

$$\left\|\begin{bmatrix}1\\2\\1\\0\end{bmatrix}\right\| = \sqrt{(1)^2+(2)^2+(1)^2+(0)^2} = \sqrt{6}$$

$$\left\|\begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}\right\| = \sqrt{(-\frac{1}{2})^2+(0)^2+(\frac{1}{2})^2+(1)^2} = \sqrt{\frac{3}{2}}$$

$$\left\|\begin{bmatrix}\frac{1}{3}\\0\\-\frac{1}{3}\\\frac{1}{3}\end{bmatrix}\right\| = \sqrt{(\frac{1}{3})^2+(0)^2+(-\frac{1}{3})^2+(\frac{1}{3})^2} = \sqrt{\frac{1}{3}}$$

On divise chaque vecteur par sa norme pour obtenir :

$$\begin{bmatrix}1\\2\\1\\0\end{bmatrix}\frac{1}{\sqrt{6}} = \begin{bmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\0\end{bmatrix} \qquad \begin{bmatrix}-\frac{1}{2}\\0\\\frac{1}{2}\\1\end{bmatrix}\sqrt{\frac{2}{3}} = \begin{bmatrix}-\frac{1}{\sqrt{6}}\\0\\\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix} \qquad \begin{bmatrix}\frac{1}{3}\\0\\-\frac{1}{3}\\\frac{1}{3}\end{bmatrix}\sqrt{3} = \begin{bmatrix}\frac{1}{\sqrt{3}}\\0\\-\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{bmatrix}$$

On a donc une base orthonormale pour U :

$$\left\{\begin{bmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\0\end{bmatrix}, \begin{bmatrix}-\frac{1}{\sqrt{6}}\\0\\\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix}, \begin{bmatrix}\frac{1}{\sqrt{3}}\\0\\-\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{bmatrix}\right\}$$

Exercice: Transformer les suivants en bases orthonormales, selon Gram-Schmidt.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\-2\\0\end{bmatrix}, \begin{bmatrix}0\\0\\,\sqrt{2}\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\-1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\2\\3\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}13\\-9\end{bmatrix}\right\}$$

Exercice: Par un argument géométrique, montrer qu'il y a exactement deux bases orthonormaux pour $\mathbb{R}^2$ tel que un des vecteurs dans la base est $\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}$. Donc, transformer la base $\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}a\\b\end{bmatrix}\right\}$ en base orthonormale selon Gram-Schmidt.

1. Une Invariante d'une Matrice

On se souvient d'un règle spécial pour calculer l'inverse d'une matrice de taille 2×2 :

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

On peut vérifier facilement en multipliant ! Ce qui est intéressant c'est que cette formule est équipée avec une condition : elle ne peut s'appliquer que si $ad - bc \neq 0$. On a vu aussi que l'inverse existe que si $ad - bc \neq 0$. C'est une invariante numérique de la matrice qui accomplit deux choses : elle indique si la matrice est inversible, et elle forme le début du calcul de l'inverse.

Définition 1. *Soit* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *une matrice* 2×2.

Alors le DÉTERMINANT *de* A *est* $ad - bc$. *On écrit soit* $\det(A)$ *ou* $|A|$ *pour le déterminant.*

On s'inspire de cette définition pour comprendre les déterminants en générale. Un COFACTEUR d'une matrice est la matrice qu'on obtient en enlevant une rangée et une colonne. Précisément, on écrit $C_{ij}(A)$ pour la matrice qu'on obtient de A en enlevant la rangée i et la colonne j.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}, \quad C_{2,3}(A) = (-1)^5 \begin{bmatrix} 1 & 2 & 4 \\ 9 & 10 & 12 \\ 13 & 14 & 16 \end{bmatrix} = -1 \begin{bmatrix} 1 & 2 & 4 \\ 9 & 10 & 12 \\ 13 & 14 & 16 \end{bmatrix}$$

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad C_{2,2}(B) = (-1)^4 \begin{bmatrix} d \end{bmatrix} = d$$

Définition 2. *Soit* A *une matrice* $n \times n$.

Alors $\det(A) = a_{1,1}C_{1,1}(A) + a_{1,2}C_{1,2}(A) + \cdots + a_{1,n}C_{1,n}(A)$, *où* a_{ij} *est l'élément de* A *à la position* (i,j) *et* $C_{ij}(A)$ *est le cofacteur de* A *à la position* (i,j).

On dit que c'est l'expansion par cofacteurs *sur la première rangée.*

Exemple: Voici le calcul d'un déterminant par expansion.

$$\det\begin{bmatrix}1&2&3&4\\0&2&-1&0\\1&2&3&4\\4&5&0&-1\end{bmatrix} = (1)(-1)^{1+1}\det\begin{bmatrix}2&-1&0\\2&3&4\\5&0&-1\end{bmatrix} + (2)(-1)^{1+2}\det\begin{bmatrix}0&-1&0\\1&3&4\\4&0&-1\end{bmatrix}$$
$$+ (3)(-1)^{1+3}\det\begin{bmatrix}0&2&0\\1&2&4\\4&5&-1\end{bmatrix} + (4)(-1)^{1+4}\det\begin{bmatrix}0&2&-1\\1&2&3\\4&5&0\end{bmatrix}$$

On voit que calculer un déterminant 4×4 nécessite calculer 4 déterminants 3×3.

$$\begin{aligned}\det\begin{bmatrix}2&-1&0\\2&3&4\\5&0&-1\end{bmatrix} &= (2)(-1)^{1+1}\det\begin{bmatrix}3&4\\0&-1\end{bmatrix} + (-1)(-1)^{1+3}\det\begin{bmatrix}2&4\\5&-1\end{bmatrix}\\ &\quad + (0)(-1)^{1+3}\det\begin{bmatrix}2&3\\5&0\end{bmatrix}\\ &= (2)\big[(3)(-1)-(4)(0)\big] - (-1)\big[(2)(-1)-(4)(5)\big] + (0)\big[(2)(0)-(3)(5)\big]\\ &= -28\end{aligned}$$

On voit que le terms $(-1)^{i+j}$ contribue seulement une alternance de signes. Aussi, on peut ignorer des termes correspondants à $a_{ij}=0$.

$$\begin{aligned}\det\begin{bmatrix}0&-1&0\\1&3&4\\4&0&-1\end{bmatrix} &= -(-1)\det\begin{bmatrix}1&4\\4&-1\end{bmatrix}\\ &= -(-1)\big[(1)(-1)-(4)(4)\big] = -17\\ \det\begin{bmatrix}0&2&0\\1&2&4\\4&5&-1\end{bmatrix} &= -(2)\det\begin{bmatrix}1&4\\4&-1\end{bmatrix}\\ &= -(2)\big[(1)(-1)-(4)(4)\big] = 34\\ \det\begin{bmatrix}0&2&-1\\1&2&3\\4&5&0\end{bmatrix} &= -(2)\det\begin{bmatrix}1&3\\4&0\end{bmatrix} + (-1)\det\begin{bmatrix}1&2\\4&5\end{bmatrix}\\ &= -(2)\big[(1)(0)-(3)(4)\big] + (-1)\big[(1)(5)-(2)(4)\big] = 27\end{aligned}$$

Finalement on met le tout ensemble pour avoir :

$$\det\begin{bmatrix}1&2&3&4\\0&2&-1&0\\1&2&3&4\\4&5&0&-1\end{bmatrix} = (1)\big[-28\big] - (2)\big[-17\big] + (3)(-1)^{1+3}\big[34\big] - (4)(-1)^{1+4}\big[27\big] = 0$$

Calculer le déterminant d'une matrice $n \times n$ nécessite le calcul de n déterminants de taille $n-1$, chacun nécessitant $n-1$ déterminants de taille $n-2$. En total, on calcul $n!/2$ déterminants de tailles 2×2.[1] C'est long ! On verra une autre méthode...

La première simplification est qu'on n'a pas besoin de prendre la première rangée.

Théorème 3. *On peut calculer le déterminant en faisant une expansion par cofacteurs pour n'importe quelle rangée ou colonne.*

Le grand avantage c'est qu'on peut choisir la rangée ayant le plus de zéros possible. Ici, on choisit de faire une expansion par la deuxième colonne, ensuite par la troisième colonne de la matrice 3×3.

Proposition 4. *Soit A une matrice $n \times n$. Alors* $\det(A) = \det(A^T)$.

Exemple:

$$\det \begin{bmatrix} 4 & 0 & 3 & 5 \\ 1 & 0 & 2 & 0 \\ 0 & 6 & 0 & 7 \\ 9 & 0 & 8 & 0 \end{bmatrix} = -(6) \det \begin{bmatrix} 4 & 3 & 5 \\ 1 & 2 & 0 \\ 9 & 8 & 0 \end{bmatrix}$$

$$= -6 \cdot (5) \det \begin{bmatrix} 1 & 2 \\ 9 & 8 \end{bmatrix} = -30 \cdot \big[(1)(8) - (9)(2)\big] = 300$$

2. Matrices Triangulaires

Une matrice TRIANGULAIRE SUPÉRIEURE est une matrice A tel que $A_{ij} = 0$ si $i < j$. C'est plus clair en dessin : la partie en bas du diagonal est zéro, donc la matrice a la forme d'un triangle. Une matrice TRIANGULAIRE INFÉRIEURE est zéro en haut du digaonal. Voici une matrice triangulaire supérieure, une matrice triangulaire inférieure, et une matrice qui est triangulaire dans les deux sens :

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

En générale, calculer le déterminant peut entraîner un grand nombre d'opérations. mais ce n'est pas le cas pour une matrice triangulaire.

Théorème 5. *Soit A une matrice triangulaire (inférieure ou supérieure).*

Alors $\det(A)$ *est le produit des éléments sur le diagonal.*

[1] Le factoriel, $n!$, est le produit des entiers de 1 à n ; donc $n! = n(n-1)\cdots(2)(1)$.

La démonstration repose sur un choix assidu d'expansion : on choisit toujours la rangée (ou colonne) ayant le plus de zéros. La diagonalité garantit qu'il il y une rangée (et une colonne) ayant qu'un seul non-zéro. $\square$

Exemple: Ici, on fait l'expansion toujours par la première colonne :

$$\begin{aligned}\det\begin{bmatrix}1 & -2 & 0 & 3\\ 0 & 4 & 0 & 2\\ 0 & 0 & -1 & 2\\ 0 & 0 & 0 & 8\end{bmatrix} &= (1)\det\begin{bmatrix}4 & 0 & 2\\ 0 & -1 & 2\\ 0 & 0 & 8\end{bmatrix}\\ &= (1)(4)\det\begin{bmatrix}-1 & 2\\ 0 & 8\end{bmatrix}\\ &= (1)(4)\big[(-1)(8)-(2)(0)\big] = (1)(4)(-1)(2) = -8\end{aligned}$$

Note que c'est les zéros en-dessous de la diagonale qui sont importants : en haut on peut avoir n'importe quoi, incluant des zéros. On découvre un résultat. Si une matrice est A triangulaire, alors $\det(A) = 0$ exactement lorsque A possède moins que n pivots. Donc A est inversible exactement lorsque $\det(A) \neq 0$. On a déja vu ceci pour les matrices 2×2 ; maintenant on connaît que c'est vrai pour toute matrice triangulaire (donc pour toute matrice en forme échelonnée !). Dans la prochaine section on verra que c'est vrai pour toute matrice, et on découvrera une méthode plus efficace de calculer les déterminants.

3. Déterminants et Réduction

Les opérations de rangée sont utiles afin de résoudre des systèmes linéaires. On a vu aussi qu'elles conservent l'espace rangée et le noyau. On veut savoir maintenant leur effet sur les déterminants

Théorème 6. *Soit A une matrice.*

I. Si B est obtenu de A par l'opération $R_i \rightleftharpoons R_j$, alors $\det(B) = -\det(A)$.

II. Si B est obtenu de A par l'opération $R_i \to tR_i$, alors $\det(B) = t\det(A)$.

III. Si B est obtenu de A par l'opération $R_i \to R_i + tR_j$, alors $\det(B) = \det(A)$.

La preuve de ce théorème découle de la théorie des matrices élémentaires. Chaque opération de rangée correspond à une multiplication par une matrice élémentaire, et on peut déduire exactement l'effet sur le déterminant. On n'a pas le temps d'en parler dans ce cours !

En termes pratique, on peut réduire une matrice à une forme échelonnée, et ensuite calculer l'inverse de la version réduite, qui est diagonale.

Recalculons les déterminants ci-haut en utilisant une réduction.

Exemple:

$$\begin{bmatrix} 4 & 0 & 3 & 5 \\ 1 & 0 & 2 & 0 \\ 0 & 6 & 0 & 7 \\ 9 & 0 & 8 & 0 \end{bmatrix} \xrightarrow[R_1 \rightleftharpoons R_2]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 4 & 0 & 3 & 5 \\ 0 & 6 & 0 & 7 \\ 9 & 0 & 8 & 0 \end{bmatrix} \xrightarrow[R_4 \to R_4 - 9R_1]{R_2 \to R_2 - 4R_1} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & -5 & 5 \\ 0 & 6 & 0 & 7 \\ 0 & 0 & -10 & 0 \end{bmatrix}$$

$$\xrightarrow[R_2 \rightleftharpoons R_3]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 6 & 0 & 7 \\ 0 & 0 & -5 & 5 \\ 0 & 0 & -10 & 0 \end{bmatrix} \xrightarrow[R_3 \to -\frac{1}{5} R_3]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 6 & 0 & 7 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -10 & 0 \end{bmatrix}$$

$$\xrightarrow[R_4 \to R_4 + 10R_3]{} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 6 & 0 & 7 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -10 \end{bmatrix}$$

Posons $A = \begin{bmatrix} 4 & 0 & 3 & 5 \\ 1 & 0 & 2 & 0 \\ 0 & 6 & 0 & 7 \\ 9 & 0 & 8 & 0 \end{bmatrix}$ et $B = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 6 & 0 & 7 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -10 \end{bmatrix}$.

Alors on a $\det(B) = (-1)(-1)(-\frac{1}{5}) \det(A)$. Mais puisque B est triangulaire, on connait que $\det(B) = (1)(6)(1)(-10) = -60$. Donc $\det(A) = -5 \det(B) = 300$.

Exemple:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & -1 & 0 \\ 1 & 2 & 3 & 4 \\ 4 & 5 & 0 & -1 \end{bmatrix} \xrightarrow[R_3 \to R_3 - R_1]{} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 4 & 5 & 0 & -1 \end{bmatrix}$$

Le déterminant de la première matrice est égale au déterminant de la deuxième, car cette opération de rangée ne change pas le déterminant. Mais le déterminant de la deuxième est zéro : considérer l'expansion par la troisième rangée.

On découvre le règle suivant :

Proposition 7. *Si A possède une rangée nulle, alors* $\det(A) = 0$.
Si A possède deux rangées identiques, alors $\det(A) = 0$.
Si les rangées de A sont dépendantes, alors $\det(A) = 0$.

Pour la première, on fait une expansion par cette rangée. Pour la deuxième, on soustrait les deux rangées, créeant aisni une rangée nulle. Pour la troisième, on pourra éventuellement obtenir une rangée nulle en réduisant. □

Théorème 8. *Soit A une matrice $n \times n$.*

(1) $\det(tA) = t^n \det(A)$

(2) $\det(AB) = \det(A)\det(B)$

(3) $\det(A) = 0$ *si et seulement si A n'est pas inversible.*

(4) Si A est inversible, alors $\det(A^{-1}) = \frac{1}{\det(A)}$

Pour la première, on observe que multiplier A par t est la même chose que multiplier chaque rangée de A par t. Pour chaque multiplication de rangée, on a un facteur de t. La deuxième découle de la théorie des matrice élémentaires (on n'en parlera pas). La troisième se voit en considérant le dioagonal dans la forme échelonnée de A. La quatrième est une conséquence du règle des produits : $\det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$. Donc si est est inversible, $\det(A^{-1}) = \det(I)/\det(A) = 1/\det(A)$. □

On a vu beaucoup de conditions équivalentes. Voici un résumé :

Théorème 9. *Soit A une matrice $n \times n$. Alors toutes les conditions suivantes sont équivalentes : elles sont soit toutes vraies ou soit toutes fausses.*

(1) $\text{rang}(A) = n$

(2) $\det(A) \neq 0$

(3) A est inversible

(4) A possède un pivot dans chaque rangée

(5) Les rangées de A sont indépendantes

(6) Les colonnes de A engendrent $\mathbb{R}^n$

(7) $A\mathbf{x} = \mathbf{b}$ *est consistant pour tout* $\mathbf{b}$

(8) A possède un pivot dans chaque colonne

(9) Les colonnes de A sont indépendantes

(10) Les rangées de A engendrent $\mathbb{R}^n$

(11) $A\mathbf{x} = \mathbf{0}$ *possède une solution unique*

(12) $A\mathbf{x} = \mathbf{b}$ *possède une solution unique pour tout* $\mathbf{b}$

(13) 0 *n'est pas valeur propre de A*

C'est une spécialisation des Quatre Théorèmes **pour des matrices carrées**. En générale, ce n'est pas vrai que l'indépendance des colonnes est équivalent à l'indépendance des rangées. Valeur propre ? On verra dans le prochain chapitre…

MAT 1741	**22. Valeurs Propres, Vecteurs Propres, Espaces Propres**	
Mike Newman	Université d'Ottawa	automne 2008

1. Motivation

En solutionnant des systèmes linéaires, on a découvert le rang d'une matrice. C'est un concept très important. Sachant le rang, on connaît le genre de solutions (unique, infini) et leur existence ; on connaît l'inversabilité de la matrice. On connaît les dimensions des quatre espaces. Tout ceci découle de la théorie des équations matricielles $A\mathbf{x} = \mathbf{b}$.

Par contre, il y a beaucoup d'applications ou on ne cherche pas a résoudre un système $A\mathbf{x} = \mathbf{b}$, mais on cherche plutôt à résoudre un système $A\mathbf{x} = \lambda\mathbf{x}$, ou $\mathbf{x}$ est un vecteur inconnu et λ est un chiffre inconnu. C'est une équation de VALEUR ET VECTEUR PROPRE.

En voivi quelques exemples. Ce serait peut-être utile de relire ces exemples *après* que vous connaissez un peu plus. Pour le moment, c'est pour donner une idée seuelement. Le cours comme soit continue à la prochaine section.

Exemple: Imaginons qu'on cherche a évaluer l'importance de différants pages web. On pourrait imaginer que si une page web P est importante, alors d'autres pages auront des liens à celle-ci. Donc : l'importance de P est proportionnelle à la somme des importances des pages ayant des liens à P. Problème : il faut connaitre toutes les importances afin de calculer les importances. . .

On construit donc une matrice (énorme !) tel que les rangées et colonnes correspondent à toutes les pages web : la première page web correspond à la première rangée et à la première colonne, etc. À la position (i, j) on met soit 1 si j possède un lien vers i, et 0 sinon. On nomme cette matrice A : elle est de taille $n \times n$, où n est le nombre de pages web. Soit $\mathbf{x}$ est un vecteur dans $\mathbb{R}^n$ où chaque x_i correspond à l'importance de la page web i.

On cherche alors $\mathbf{x}$ tel que $A\mathbf{x} = \lambda\mathbf{x}$ pour une constante λ. C'est une équation de valeur et vecteur propre.

La théorie date des années 1950, de Kendall et Wei. Aujourd'hui, Google utilise une variante de cette idée pour trier les sites web en ordre d'importance. Voir par exemple `http://www.math.upenn.edu/~wilf/website/KendallWei.pdf` pour plus de détails, ainsi que les références là-dedans.

Exemple: On se souvient de l'équation d'une ellipse :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

où le rayon dans la direction de l'axe des x est a, et le rayon dans la direction de l'axe des y est b. Mais si l'ellipse n'est pas horizontale ou verticale ?

En générale l'équation d'une ellipse est une FORME QUADRATIQUE :

$$0 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Les axes principaux sont dans la direction des vecteurs propres, et les rayons sont déterminés par les valeurs propres de la matrice $\left[\begin{smallmatrix} \alpha & \beta \\ \beta & \gamma \end{smallmatrix}\right]$.

Exemple: On cherche à modeler une maladie. Les personnes sont soit en bonne santé, infectés, ou malades. C'est trois différents états, et on a donc une matrice de transition de taille 3×3. La valeur dans la position (i, j) de cette matrice correspond à la probabilité qu'un individu en état i serait prochainement en état j. Cette matrice possède toujours la valeur propre 1, et le vecteur propre correspondant donne la situation d'équilibre. Par exemple, soit

$$A = \begin{bmatrix} .8 & .1 & .5 \\ .2 & .8 & 0 \\ 0 & .1 & .5 \end{bmatrix}$$

La première colonne détermine le sort d'un individu en bonne santé : il a 80% de chances de rester en bonne santé et 20% de chance d'être infecté. La deuxième colonne correspond aux individus infectés : %10 de chances de se guérir, 10% de chance de tomber carrément malade, et 80% de chance de rester infectés sans le savoir. La troisième colonne indique qu'un individu qui est malade à 50% de chance à se guérir et 50% de chance à rester malade. (on imagine que ces probabilités correspondent à une intervalle de temps, eg, chaque jour ou chaquie semaine).

En commençant avec x_1 personnes en bonne santé, x_2 personnes infectés et x_3 personnes malades, les valeurs pour la prochaine étape sont calculées selon $A\mathbf{x}$.

On s'intéresse à l'équilibre : est-ce que $x_3 \to 0$? Est-ce que $x_1 \to 0$? Ou est-ce qu'il y aura un point d'équilibre non-triviale ?

Le point d'équilibre correspond à $A\mathbf{x} = \mathbf{x}$: le vecteur $\mathbf{x}$ est vecteur propre correspondant à la valeur propre 1.

2. Valeurs et Vecteurs Propres

On s'intéresse à l'équation $A\mathbf{x} = \lambda\mathbf{x}$. Ici, $\mathbf{x}$ est un vecteur inconnu et λ est un chiffre inconnu.

Exercice: Montrer que l'équation $A\mathbf{x} = \lambda\mathbf{x}$ n'a aucun sens si A n'est pas carré.

Définition 1. *Soit A une matrice de taille $n \times n$.*

Si $\mathbf{x}$ est un vecteur de dimension n, $\mathbf{x} \neq \mathbf{0}$, et λ est un chiffre tel que $A\mathbf{x} = \lambda\mathbf{x}$, alors $\mathbf{x}$ est VECTEUR PROPRE *de A et λ est* VALEUR PROPRE *de A.*

En anglais, on dit "eigenvector" et "eigenvalue". Le mot "eigen" est allemand, voulant dire... propre, dans le sens de propre à soi.

Afin de vérifier si un vecteur est vecteur propre, on multiplie la matrice et le vecteur.

Exemple: Montrer que $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ est vecteur propre de $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$.

On calcul $\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Donc $A\mathbf{x} = 3\mathbf{x}$. Alors $\mathbf{x}$ est vecteur propre correspondant à la valeur propre 3.

Exemple: Montrer que $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ n'est pas vecteur propre de $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$.

On calcul $\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$. On voit que $\begin{bmatrix} 4 \\ 7 \end{bmatrix}$ n'est pas un multiple de $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Donc $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ n'est pas vecteur propre.

Exemple: Montrer que $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ n'est pas vecteur propre de $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$.

On ne calcul rien du tout : le vecteur zéro n'est *jamais* vecteur propre !

Exemple: Montrer que $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ est vecteur propre de $\begin{bmatrix} 2 & 5 & -7 \\ -1 & -4 & 5 \\ -1 & -1 & 2 \end{bmatrix}$.

On calcul $\begin{bmatrix} 2 & 5 & -7 \\ -1 & -4 & 5 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Donc $A\mathbf{x} = 0\mathbf{x}$. Alors $\mathbf{x}$ est vecteur propre correspondant à la valeur propre 0. Autrement dit, $\mathbf{x}$ est un vecteur dans le noyau de A !

Afin de vérifier si un chiffre est valeur propre, on... ?

Exemple: Montrer que 1 est valeur propre de $A = \begin{bmatrix} .8 & .1 & .5 \\ .2 & .8 & 0 \\ 0 & .1 & .5 \end{bmatrix}$

On cherche un vecteur $\mathbf{x}$ tel que $A\mathbf{x} = 1\mathbf{x}$.

Par inspiration, on essaie le vecteur $\begin{bmatrix} \frac{5}{17} \\ \frac{10}{17} \\ \frac{2}{17} \end{bmatrix}$. On voit que

$$\begin{bmatrix} .8 & .1 & .5 \\ .2 & .8 & 0 \\ 0 & .1 & .5 \end{bmatrix} \begin{bmatrix} \frac{5}{17} \\ \frac{10}{17} \\ \frac{2}{17} \end{bmatrix} = \begin{bmatrix} \frac{5}{17} \\ \frac{10}{17} \\ \frac{2}{17} \end{bmatrix} = 1 \cdot \begin{bmatrix} \frac{5}{17} \\ \frac{10}{17} \\ \frac{2}{17} \end{bmatrix}$$

Donc c'est vecteur propre, correspondant à la valeur propre 1.

Suivant l'exemple de maladie ci-haut, on a une interprétation de ce vecteur : au long terme, $\frac{5}{17}$ de la population serait en bonne santé, $\frac{10}{17}$ serait infecté mais pas malade et $\frac{2}{17}$ de la population serait malade.

3. Polynôme Caractéristique

Étant donné une matrice A, on peut facilement déterminer si un vecteur $\mathbf{x}$ est vecteur propre de A : on multiple $A\mathbf{x}$; si c'est un multiple de $\mathbf{x}$, alors $\mathbf{x}$ est vecteur propre (est le multiple est la valeur propre). Mais comment savoir si un chiffre est valeur propre ?

On cherche une solution $\mathbf{x}$ et λ pour l'équation $A\mathbf{x} = \lambda\mathbf{x}$. On peut transformer l'équation comme suit :

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ \mathbf{0} &= \lambda\mathbf{x} - A\mathbf{x} \\ \mathbf{0} &= \lambda I\mathbf{x} - A\mathbf{x} \\ \mathbf{0} &= (\lambda I - A)\mathbf{x} \end{aligned}$$

Donc si il y a une solution, alors $\mathbf{x}$ est dans $\ker(\lambda I - A)$. Aussi, $\mathbf{x} \neq \mathbf{0}$. Donc ce noyau n'est pas triviale. Alors $B = \lambda I - A$ est une matrice carrée ayant un noyau non-triviale. Donc : B n'est pas inversible, B possède moins que n pivots, les colonnes de B sont dépendantes ... et $\det(B) = 0$.

La condition $\det(B) = 0$ n'a rien à faire avec $\mathbf{x}$: c'est un polynôme de degré n dans la variable λ. Les racines de ce polynôme sont exactement les valeurs propres de A.

Définition 2. *Le polynôme caractéristique de A est* $\det(\lambda I - A)$.

Le Théorème fondamental de l'algèbre (vu dans le prmier chapitre) dit que tout polynôme de degré n ayant des coefficients réels ou complexes possède n racines, incluant leur multiplicité. Conséquence : une matrice de taille $n \times n$ possède n valeurs propres, incluant leur multiplicité.

Exemple: Trouver les valeurs et vecteurs propres de $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$.

On calcul

$$\begin{aligned} \det\left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}\right) &= \det \begin{bmatrix} \lambda - 2 & -1 \\ 1 & \lambda - 4 \end{bmatrix} \\ &= (\lambda - 2)(\lambda - 4) - (-1)(1) \\ &= \lambda^2 - 6\lambda + 9 \\ &= (\lambda - 3)^2 \end{aligned}$$

Donc il y a une seule valeur propre : $\lambda = 3$.

Les vecteurs propres correspondants sont les $\mathbf{x}$ tel que $(3I - A)\mathbf{x} = \mathbf{0}$. Donc on cherche $\ker(3I - A)$. On trouve sa forme échelonnée réduite :

$$\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

Donc une base pour $\ker(3I - A)$ est

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Les vecteurs propres sont exactement les multiples de $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.

Définition 3. *Soit une matrice carré A.*

Le POLYNÔME CARACTÉRISTIQUE *de A est* $\det(\lambda I - A)$.

Soit $\lambda = t$ une racine du polynôme caractéristique.

Sa MULTIPLICITÉ ALGÉBRIQUE *est sa multiplicité comme racine du polynôme.* *Sa* MULTIPLICITÉ GÉOMÉTRIQUE *est la dimension de* $\ker(tI - A)$.

Dans l'exemple précédent, la multiplcité algébrique de $\lambda = 3$ est deux ; sa multiplicité géométrique est un.

Note que les vecteurs propres correspondants à une valeur propre $\lambda = t$ sont exactement les veceteurs non-nuls du noyau de $tI - A$. On dit que ce noyau est un ESPACE PROPRE. La meilleur façon de "donner les vecteurs propres" c'est de donner une base pour l'espace propre.

Définition 4. *Soit A une matrice $n \times n$, et $\lambda = t$ une valeur propre.*

Alors L'ESPACE PROPRE *correspondant à t est* $\ker(tI - A)$.

Voyons un autre exemple, après lequel on énoncera la méthode générale.

Exemple: Trouver les valeurs propres de la matrice $A = \begin{bmatrix} 2 & 5 & -7 \\ -1 & -4 & 5 \\ -1 & -1 & 2 \end{bmatrix}$. Donner une base pour chaque espace propre.

On calcul

$$\begin{aligned}
\det(\lambda I - A) &= \det \begin{bmatrix} \lambda-2 & -5 & 7 \\ 1 & \lambda+4 & -5 \\ 1 & 1 & \lambda-2 \end{bmatrix} \\
&= (1)\cdot\det\begin{bmatrix} -5 & 7 \\ \lambda+4 & -5 \end{bmatrix} - (1)\cdot\det\begin{bmatrix} \lambda-2 & 7 \\ 1 & -5 \end{bmatrix} + (\lambda-2)\cdot\det\begin{bmatrix} \lambda-2 & -5 \\ 1 & \lambda+4 \end{bmatrix} \\
&= \big[(-5)(-5)-(7)(\lambda+4)\big] - \big[(\lambda-2)(-5)-(7)(1)\big] \\
&\qquad + (\lambda-2)\big[(\lambda-2)(\lambda+4)-(-5)(1)\big] \\
&= \big[-7\lambda-3\big] - \big[-5\lambda+3\big] + (\lambda-2)\big[\lambda^2+2\lambda-3\big] \\
&= \big[-7\lambda-3\big] - \big[-5\lambda+3\big] + \big[\lambda^3-7\lambda+6\big] \\
&= \lambda^3 - 9\lambda \\
&= \lambda(\lambda+3)(\lambda-3)
\end{aligned}$$

Les racines sont $\lambda = 0$, $\lambda = 3$ et $\lambda = -3$. Note qu'on a eu un peu de chance : le cubique possède ici un facteur de λ, donc on peu appliquer la formule quadratique. En générale, c'est difficile de trouver les racines d'un polynôme.[1]

Les valeurs propres sont 0, 3 et -3, chacune de multiplicité algébrique 1.

Il y a trois espaces propres.

Pour $\lambda = 0$ on obitent :

$$0I - A = \begin{bmatrix} -2 & -5 & 7 \\ 1 & 4 & -5 \\ 1 & 1 & -2 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

On vous laisse la réduction comme exercice ! Une base pour le noyau, et donc une base pour l'espace propre E_0, est

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Pour $\lambda = 3$ on obtient :

$$3I - A = \begin{bmatrix} 1 & -5 & -7 \\ 1 & 7 & -5 \\ 1 & 1 & 1 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

[1]Il existe une formule générale pour les polynômes de degré trois (c'est un peu long), et même de degré quatre (c'est très long !). Mais c'est un théorème incroyable dû a Abel et Ruffini qui dit qu'il n'existe pas de formule générale pour les polynômes de degré 5 ou plus.

Ceci donne la base pour l'espace propre E_3 :

$$\left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Pour $\lambda = -3$ on obtient :

$$-3I - A = \begin{bmatrix} -5 & -5 & -7 \\ 1 & 1 & -5 \\ 1 & 1 & -5 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Ceci donne la base pour l'espace propre E_3 :

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Pour résumer, on a les valeurs propres suivantes, avec des bases pour leurs espaces propres.

$$\lambda = 0, \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \qquad \lambda = 3, \left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right\} \qquad \lambda = -3, \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Voici un exercice très important.

Exercice: Soit $\lambda = t$ une valeur propre d'une matrice A. Montrer que la forme échelonnée réduite de $B = tI - A$ ne possède jamais un pivot dans chaque rangée.

C'est impossible d'avoir une valeur propre sans vecteur propre. C'est la définition même ! C'est aussi l'exercice précédent. Donc si, en trouvant la forme échelonnée de $tI - A$, on trouve un pivot dans chaque rangée, alors on sait qu'il y a une erreur.

Théorème 5. *La multiplicité géométrique d'une valeur propre est au moins un.*

Exercice: Trouver toutes les valeurs propres de $A = \begin{bmatrix} .8 & .1 & .5 \\ .2 & .8 & 0 \\ 0 & .1 & .5 \end{bmatrix}$. N'oublier pas qu'on connaît déja une valeur propre, donc on connaît déja un facteur du polynôme caractéristique. En principe, vous êtes capable de trouver les vecteurs propres aussi, mais les calculs sont difficiles.

Terminologie que vous devriez comprendre : valeur propre, vecteur propre, espace propre, polynôme caractéristique, multiplicité algébrique, multiplicité géométrique.

Mike Newman Université d'Ottawa automne 2008

1. Valeurs et Vecteurs

On rappelle la définition. Soit $A\mathbf{x} = \lambda\mathbf{x}$, avec $\mathbf{x} \neq \mathbf{0}$. Alors

(1) λ est valeur propre de A

(2) $\mathbf{x}$ est vecteur propre correspondant à λ

(3) E_λ est l'ensemble de tout vecteur propre correspondant à λ.

NB: C'est possible d'avoir $\lambda = 0$, même $A = \mathbf{0}$, mais jamais $\mathbf{x} = \mathbf{0}$. Les vecteurs propres ne sont jamais nuls ![1]

Exercice: Donner toutes les valeurs, vecteurs et espaces propres de $A = \mathbf{0}$.

> **Algorithme 1.** *Soit A une matrice de taille $n \times n$.*
>
> *(1) On trouve le polynôme caractéristique en évaluant $p(\lambda) = \det(\lambda I - A)$. C'est un polynôme de degré n.*
>
> *(2) Les valeurs propres sont exactement les racines de $p(\lambda)$. Pour chaque valeur propre $\lambda = r$, sa* MULTIPLICITÉ ALGÉBRIQUE *est sa multiplicité comme racine de $p(\lambda)$.*
>
> *(3) Pour chaque valeur propre $\lambda = r$, il y a un espace propre correspondant : $E_r = \ker(rI - A)$. On donne une base pour l'espace propre en donnant une base pour $\ker(rI - A)$.*
>
> *(4) Pour chaque valeur propre $\lambda = r$, sa* MULTIPLICITÉ GÉOMÉTRIQUE *est la dimension de $\ker(rI - A)$.*

On a déja vu des exemples de cet algorithme, dans le dernier chapitre et aussi en classe. Note que chaque étape (déterminant, noyau, base) et quelque chose qu'on a déja vu.

Attention! En calculant le polynôme caractéristique, il ne faut pas oublier qu'on veut les *racines* de ce polynôme. Donc comme règle générale, on ne va pas tout multiplier pour obtenir un polynôme de degré n, on va plutôt la calculer soigneusement. Si on voit un facteur commun, on la laisse en commun.

[1] Donc strictement, il y a une erreur : E_λ est l'ensemble de tout vecteur propre correspondant à λ, ainsi que le vecteur zéro. Dit d'une autre façon, E_λ est l'ensemble de toutes les solutions à $A\mathbf{x} = \lambda\mathbf{x}$, et les vecteurs propres sont les solutions non-triviales. Bref : les vecteurs propres ne sont jamais nuls, mais on permet le vecteur $\mathbf{0}$ dans E_λ. Pourquoi ? Pour que E_λ soit un sous-espace.

Exemple: Factoriser le polynôme $(\lambda-3)^2(\lambda^2+5\lambda+6)$.

Méthode A :

$$\begin{aligned}(\lambda-3)^2(\lambda^2+5\lambda+6) &= (\lambda^2-6\lambda+9)(\lambda^2+5\lambda+6)\\ &= \lambda^4-\lambda^3+4\lambda^2+9\lambda+54\\ &= \ldots ???\end{aligned}$$

Méthode B :

$$(\lambda-3)^2(\lambda^2+5\lambda+6) = (\lambda-3)^2(\lambda+2)(\lambda+3)$$

Exemple: Trouver les valeurs et vecteurs propres de $A = \begin{bmatrix} 3 & 1 & 0 & 2\\ 0 & 2 & 5 & -7\\ 0 & -1 & -4 & 5\\ 0 & -1 & -1 & 2\end{bmatrix}$.

On calcul

$$\begin{aligned}\det(\lambda I - A) &= \begin{bmatrix} \lambda-3 & -1 & 0 & -2\\ 0 & \lambda-2 & -5 & 7\\ 0 & 1 & \lambda+4 & -5\\ 0 & 1 & 1 & \lambda-2\end{bmatrix}\\ &= (\lambda-3)\det\begin{bmatrix} \lambda-2 & -5 & 7\\ 1 & \lambda+4 & -5\\ 1 & 1 & \lambda-2\end{bmatrix}\\ &= (\lambda-3)\lambda(\lambda+3)(\lambda-3)\\ &= \lambda(\lambda+3)(\lambda-3)^2\end{aligned}$$

On a de la chance : le déterminant 3×3 était déja calculé dans le chapitre précédent.

Les valeurs propres sont 0, -3 et $+3$, de multiplicité algébrique un, un et deux. On résume les caluls des bases pour chaque espace propre (vérifier les détails !)

Pour $\lambda = 0$:

$$0I - A = \begin{bmatrix} -3 & -1 & 0 & -2\\ 0 & -2 & -5 & 7\\ 0 & 1 & 4 & -5\\ 0 & 1 & 1 & -2\end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 0 & 1\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & -1\\ 0 & 0 & 0 & 0\end{bmatrix} \qquad \text{base : } \left\{\begin{bmatrix} -1\\ 1\\ 1\\ 1\end{bmatrix}\right\}$$

Pour $\lambda = -3$:

$$-3I - A = \begin{bmatrix} -6 & -1 & 0 & -2\\ 0 & -5 & -5 & 7\\ 0 & 1 & 1 & -5\\ 0 & 1 & 1 & -5\end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & -\frac{1}{6} & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix} \qquad \text{base : } \left\{\begin{bmatrix} 1\\ -6\\ 6\\ 0\end{bmatrix}\right\}$$

Pour $\lambda = 3$:

$$3I - A = \begin{bmatrix} 0 & -1 & 0 & -2 \\ 0 & 1 & -5 & 7 \\ 0 & 1 & 7 & -5 \\ 0 & 1 & 1 & 1 \end{bmatrix} \to \cdots \to \begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{base : } \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Note que chaque fois, la multiplicité géomt́rique (dimension de $\ker(rI - A)$) est le nombre de colonnes sans pivot. C'est comme normal ! Puisqu'il y a toujours un vecteur propre pour chaque valeur propre, la multiplicité géomt́rique est toujours au moins un, et don il y aurau toujours au moins une colonne sans pivot.

2. Bases

Les vecteurs propres sont utiles pour plusieurs raisons, mais on peut les comprender de façon intuitive : ce sont les directions pour laquelle la matrice "agit" comme un chiffre. Les vecteurs propres sont les directions simples *par rapport à cette matrice.* On connaît l'utilité des bases pour un espace vectoriel. Maintenant on voudrait une base simple pour une application précise : *simple par rapport à une matrice donnée.*

Le premier résultat est que si on prend les bases pour chaque espace propre et on les met tous ensemble, le tout est indépendant.

Théorème 2. *Soit A une matrice de taille $n \times n$, ayant valeurs propres distinctes $\lambda_1, \lambda_2, \ldots, \lambda_m$, où $m \leq n$.*

Soit $\mathcal{B}_i$ une base pour l'espace propre E_{λ_i}.

Alors l'ensemble de toutes les $\mathcal{B}_{\lambda_i}$ est indépendant.

Chaque base propre est certainement indépendant : c'est une base ! Ce qui est surprenant, c'est l'indépendance de *toutes* ces bases mises ensemble.

Exemple: Pour la matrice $A = \begin{bmatrix} 3 & 1 & 0 & 2 \\ 0 & 2 & 5 & -7 \\ 0 & -1 & -4 & 5 \\ 0 & -1 & -1 & 2 \end{bmatrix}$ on a trouvé les bases pour les trois espaces propres :

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\} \qquad \left\{ \begin{bmatrix} 1 \\ -6 \\ 6 \\ 0 \end{bmatrix} \right\} \qquad \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Attention! Il ne s'agit pas de trois bases pour $\mathbb{R}^4$. Il s'agit pour trois bases, chacune pour un sous-espace différent. Les trois sous-espaces sont exactement les espaces propres de A.

Si on met ces trois bases ensemble on obtient :

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -6 \\ 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Le théorème garantit que c'est un ensemble de quatre vecteurs indépendants. Mais étant un ensemble de quatre vecteurs indépendants dans $\mathbb{R}^4$, c'est nécessairement une base pour $\mathbb{R}^4$. En générale, pour une matrice A de taille $n \times n$, on aurait toujours des vecteurs indépendants ; si on a la chance d'avoir exactement n vecteurs, ce serait alors automatiquement une base pour $\mathbb{R}^n$.

3. Multiplicités

Comment savoir si l'ensemble des bases pour les espaces propres a la bonne taille ?

Théorème 3. *Soit λ une valeur propre de A.*

Alors la multiplicité géométrique de λ est au moins un et au plus la multiplicité algébrique de λ.

La somme de toutes les multiplicités algébriques est exactement le degré du polynôme caractéristique, ce qui est n, la taille de A. Par contre si on veut que l'ensemble de toutes les bases propres de A soit une base pour $\mathbb{R}^n$, il faudrait avoir n vecteurs. Donc on a :

Théorème 4. *Soit A une matrice de taille $n \times n$.*

L'ensemble des bases pour touts les espaces propres est une base pour $\mathbb{R}^n$ si et seulement si la multiplicité algébrique est égale à la multiplicité géométrique pour chaque valeur propre.

Dans ce cas on dit que A est DIAGONALISABLE.

On verra pourquoi le mot “diagonalisable” dans un instant.

Exemple: $A = \begin{bmatrix} 3 & 1 & 0 & 2 \\ 0 & 2 & 5 & -7 \\ 0 & -1 & -4 & 5 \\ 0 & -1 & -1 & 2 \end{bmatrix}$ est diagonalisable. On a déja calculé les valeurs propres et les bases pour chaque espace propre. On a vu que la multiplicité algébrique est égale à la multiplicité géométrique pour chaque valeur propre. Donc la matrice est diagonalisable : on connaît déja une base pour $\mathbb{R}^4$ formée de vecteurs propres de A : c'est l'exemple au fin de la section précédente.

Exercice: Vérifier que cet ensemble est vraiment indépendant et qu'il engendre vraiment $\mathbb{R}^4$.

Exemple: $C = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$ n'est pas diagonalisable. On a déja calculé que la seule valeur propre est $\lambda = 3$ (chapitre précédent). Sa multiplicité algébrique est égale à deux et sa multiplicité géométrique est égale à un. N'étant pas égaux, il n'existe pas une base pour $\mathbb{R}^2$ formée de valeurs propres de C.

Dans ces exemples il a fallu calculer les bases pour chaque espace propre avant de pouvoir conclure s'il existe ou non une base pour $\mathbb{R}^n$ formée de vecteurs propres. Parfois ce n'est pas le cas. Par exemple, si la multiplicité algébrique est toujours un, l'égalité est forcée, car on a le théorème ci-haut :

$$1 \leq \text{multiplicité géométrique} \leq \text{multiplicité algébrique}$$

Théorème 5. *Soit A une matrice de taille $n \times n$.*

Si la multiplicité algébrique de chaque valeur propre est un, alors A est diagonalisable.

Exemple: La matrice $A = \begin{bmatrix} 2 & 5 & -7 \\ -1 & -4 & 5 \\ -1 & -1 & 2 \end{bmatrix}$ est diagonalisable. On a vu au chapitre précédent que les valeurs propres sont 0, 3 et -3, chacune de multiplicité algébrique un. Donc, *avant même de calculer des vecteurs propres* on sait qu'on aurait à la fin une base pour $\mathbb{R}^3$ formée de vecteurs propres. Bien sûr, il faudrait les calculer pour savoir ce qu'ils sont ! Les voici, recopiés du précédent, une base pour $\mathbb{R}^3$ formée de vecteurs propres pour A :

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Exercice: Vérifier que cet ensemble est indṕendant, et qu'il engendre $\mathbb{R}^3$. Vérifier aussi que chaque vecteur est vecteur propre en multipliant par A.

4. Diagonalisabilité

Soit A une matrice diagonalisable. Il existe donc n vecteurs propres de A qui sont linéairement indépendants : on les obtient en calculant une base pour chaque espace propre, et ensuite mettant toutes ces bases ensembles.

Posons P comme une matrice ayant ces vecteurs comme colonnes, et D une matrice diagonal ayant les valeurs propres sur le diagonal. Donc D est une matrice zéro, sauf sur le diagonal, ou se trouve les valeurs propres, répétées selon leur multiplicité.[2] On s'organise pour avoir les vecteurs et valeurs en ordre correspondant : la première valeur propre correspond au premier vecteur propre, etc. Alors on voit que $AP = PD$. Un dessin est peut-être utile ici, montrant les colonnes de P : calculer chaque colonne de AP et de PD.

On a donc que $A = PDP^{-1}$, ou $D = P^{-1}AP$.

Exemple: Soit A diagonalisable. Calculer A^3.

On imagine qu'on a déja calculé les valeurs propres, ainsi qu'une base pour $\mathbb{R}^n$ composée de vecteurs propres pour A. On a donc $A = PDP^{-1}$, ce qui permet de calculer :

$$\begin{aligned} A^3 &= (PDP^{-1})(PDP^{-1})(PDP^{-1}) \\ &= PDP^{-1}PDP^{-1}PDP^{-1} \\ &= PDIDIDP^{-1} \\ &= PD^3P^{-1} \end{aligned}$$

Exercice: Donner une formule pour A^{1000} ainsi que pour A^{-1}

Exercice: Imaginons que A est diagonalisable, et de plus que la base est *orthonormale*. On a donc $A = PDP^{-1}$. Expliquer pourquoi le calcul de P^{-1} est très facile dans ce cas.

Si il existe une base pour $\mathbb{R}^n$ formée de vecteurs propres de A, c'est la meilleure base *par rapport à* A. Pourquoi ?

Soit A une matrice, et $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ une base pour $\mathbb{R}^n$ formée de vecteurs propres de A. Donc $A\mathbf{b}_i = \lambda_i \mathbf{b}_i$ pour $i = 1, 2, \ldots, n$. Si $\mathbf{x} \in \mathbb{R}^n$, alors il existe des chiffres uniques $a_1, a_2, \ldots, a_n$ tel que

$$\mathbf{x} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_n\mathbf{b}_n$$

et donc pour calculer $A\mathbf{x}$ on calcul

$$\begin{aligned} A\mathbf{x} = A(a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_n\mathbf{b}_n) &= a_1A\mathbf{b}_1 + a_2A\mathbf{b}_2 + \cdots + a_nA\mathbf{b}_n \\ &= a_1\lambda_1\mathbf{b}_1 + a_1\lambda_2\mathbf{b}_2 + \cdots + a_n\lambda_n\mathbf{b}_n \end{aligned}$$

Dans la dernière expression, la multiplication est scalaire fois vecteur, au lieu de matrice fois vecteur.

Si A représente une transformation linéaire (notre prochain sujet), alors c'est plus facile de comprendre l'action de A sur une base : une base formée de vecteurs propres de A.

Terminologie que vous devriez comprendre : multiplicité algébrique, multiplicité géométrique, diagonalisable.

[2]Multiplicité algébrique ou géométrique ? Aucune différence, la matrice est diagonalisable !

1. Transformations

Une transformation linéaire est une fonction définit sur un espace vectoriel, qui possède des propriétés spéciales.

Exemple: Soit $T(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{x}$. Alors T est une fonction, mais une fonction qui prend comme ingrédient non un chiffre mais un vecteur $\mathbf{x}$. Les ingrédients valides sont les vecteurs dans $\mathbb{R}^2$; les réponses possibles sont aussi les vecteurs dans $\mathbb{R}^2$. Note qu'on peut aussi décrire T : c'est la transformation qui échange les deux coordonnés de $\mathbf{x}$. On écrite typiquement

$$T : \mathbb{R}^2 \to \mathbb{R}^2 \qquad \text{tel que} \qquad T(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{x}$$

Donc T est une transformation *de* $\mathbb{R}^2$ *vers* $\mathbb{R}^2$. En générale, on peut avoir des espaces différents. Mais plus important, pour être linéaire, il faut satisfaire deux conditions.

Définition 1. *Soit U et V des espaces vectoriels.*
Alors $T : U \to V$ est une TRANSFORMATION LINÉAIRE *si*

(1) $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ *pour tout* $\mathbf{x}, \mathbf{y} \in U$

(2) $T(a\mathbf{x}) = aT(\mathbf{x})$ *pour tout* $\mathbf{x} \in U$ *et* $a \in \mathbb{R}$

On aura presque toujours $U = \mathbb{R}^n$ et $V = \mathbb{R}^m$ pour des entiers n et m.

Exemple: Soit $T : \mathbb{R}^2 \to \mathbb{R}^3$ défini par

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \\ x_2 - x_1 \end{bmatrix}$$

Vérifier que T est linéaire.

C'est certainement une transformation entre deux espaces vectoriels. On vérifie les deux conditions. Premièrement, la condition d'addition de vecteurs :

$$T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}+\begin{bmatrix}y_1\\y_2\end{bmatrix}\right)=T\left(\begin{bmatrix}x_1+y_1\\x_2+y_2\end{bmatrix}\right)=\begin{bmatrix}(x_1+y_1)-(x_2+y_2)\\(x_1+y_1)+(x_2+y_2)\\(x_2+y_2)-(x_1+y_1)\end{bmatrix}$$
$$=\begin{bmatrix}x_1-x_2+y_1-y_2\\x_1+x_2+y_1+y_2\\x_2-x_1+y_2-y_1\end{bmatrix}$$

$$T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)+T\left(\begin{bmatrix}y_1\\y_2\end{bmatrix}\right)=\begin{bmatrix}x_1-x_2\\x_1+x_2\\x_2-x_1\end{bmatrix}+\begin{bmatrix}y_1-y_2\\y_1+y_2\\y_2-y_1\end{bmatrix}$$
$$=\begin{bmatrix}x_1-x_2+y_1-y_2\\x_1+x_2+y_1+y_2\\x_2-x_1+y_2-y_1\end{bmatrix}$$

Les deux sont égales, donc la première condition est satisfaite. Deuxièment la condition de multiplication scalaire :

$$T\left(a\cdot\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=T\left(\begin{bmatrix}ax_1\\ax_2\end{bmatrix}\right)=\begin{bmatrix}ax_1-ax_2\\ax_1+ax_2\\ax_2-ax_1\end{bmatrix}$$

$$a\cdot T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=a\cdot\begin{bmatrix}x_1-x_2\\x_1+x_2\\x_2-x_1\end{bmatrix}=\begin{bmatrix}a(x_1-x_2)\\a(x_1+x_2)\\a(x_2-x_1)\end{bmatrix}=\begin{bmatrix}ax_1-ax_2\\ax_1+ax_2\\ax_2-ax_1\end{bmatrix}$$

Les deux sont égales, donc la deuxième condition est satisfaite.

Exemple: Montrer que $T:\mathbb{R}^2\to\mathbb{R}^2$ défini par $T(\mathbf{x})=\begin{bmatrix}0&1\\1&0\end{bmatrix}\mathbf{x}$ est une transformation linéaire.

On peut calculer avec des vecteurs $\begin{bmatrix}x_1\\x_2\end{bmatrix}$ comme ci-haut, mais on peut faire mieux, en utilisant les propriétés de la multiplication matricielle. Soit $A=\begin{bmatrix}0&1\\1&0\end{bmatrix}\mathbf{x}$.

$$T(\mathbf{x}+\mathbf{y})=A(\mathbf{x}+\mathbf{y})=A\mathbf{x}+A\mathbf{y}=T(\mathbf{x})+T(\mathbf{y})$$
$$T(a\mathbf{x})=A(a\mathbf{x})=aA\mathbf{x}=aT(\mathbf{x})$$

Exemple: Est-ce que $N:\mathbb{R}^2\to\mathbb{R}$ défini par $N\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\sqrt{x^2+y^2}$ est linéaire ?

Il suffit de trouver un contre-exemple a l'une ou l'autre des conditions. En voici :

$$N\left(\begin{bmatrix}0\\1\end{bmatrix}+\begin{bmatrix}0\\-1\end{bmatrix}\right)=N\left(\begin{bmatrix}0\\0\end{bmatrix}\right)=0$$
$$N\left(\begin{bmatrix}0\\1\end{bmatrix}\right)+N\left(\begin{bmatrix}0\\-1\end{bmatrix}\right)=1+1=2$$

Ce n'est pas une transformation linéaire. (Il y a beaucoup d'autres contre-exemples.)

Afin de déterminer si une transformation est linéaire, il faut se référer à la définition de la transformation et aux conditions de linéairité. Donc chaque fois, on a montré soit que la formule donnée pour la transformation satisfait les deux conditions, ou on a trouvé un contre-exemple. Il existe d'autres méthodes (voir des cours plus avancés, ou essayer d'en déduire vous-même), mais on n'a pas le temps de développer plus dans ce cours.

2. Matrices

Soit $T : U \to V$ une transformation linéaire, avec $\dim(U) = n$ et $\dim(V) = m$. Fixons une base pour U, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$.

Si on prend n'importe quel vecteur $\mathbf{x} \in U$, alors on peut l'écrire uniquement comme combinaison linéaire de la base : $\mathbf{x} = a_1 + a_2 + \cdots + a_{\mathbf{u}}n$. Donc

$$T(\mathbf{x}) = T\left(a_1 + a_2 + \cdots + a_{\mathbf{u}}n\right) = a_1T(\mathbf{u}_1) + a_2T(\mathbf{u}_2) + \cdots + a_nT(\mathbf{u}_n)$$

Sachant l'effet de T sur une base, son effet sur $\mathbf{x}$ est déterminé. Précisément, $T(\mathbf{x})$ est une combinaison linéaire des vecteurs $T(\mathbf{u}_1)$, $T(\mathbf{u}_2)$,. . . ,$T(\mathbf{u}_n)$.

Retournons à l'exemple ci-haut, $T : \mathbb{R}^2 \to \mathbb{R}^3$ défini par $T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right)=\begin{bmatrix}x_1-x_2\\x_1+x_2\\x_2-x_1\end{bmatrix}$

On a la base standard pour $\mathbb{R}^2$: $\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\}$. On calcul alors :

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}1-0\\1+0\\0-1\end{bmatrix}=\begin{bmatrix}1\\1\\-1\end{bmatrix}\qquad T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}0-1\\1+0\\1-0\end{bmatrix}=\begin{bmatrix}-1\\1\\1\end{bmatrix}$$

On peut maintenant calculer T de n'importe quel vecteur :

$$\begin{aligned} T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) &= T\left(a\begin{bmatrix} 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= aT\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + bT\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \\ &= a\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + b\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} \end{aligned}$$

On a prouvé que T est "multiplication par une matrice fixe". C'est la MATRICE STANDARD de la transformation.

Algorithme 2. *Soit $T : \mathbb{R}^n \to \mathbb{R}^n$.*

Afin de calculer la MATRICE STANDARD *de la transformation, on commence avec la base standard de $\mathbb{R}^n$. On calcul T sur chaque élément de la base. Ce sont les colonnes de la matrice standard.*

Exercice: Soit $S : \mathbb{R}^4 \to \mathbb{R}^2$ défini par $S\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_3 - x_4 \end{bmatrix}$. Montrer que c'est une transformation linéaire. (Il faut montrer que les deux conditions sont satisfaites : calculer la côté gauche ete droite pour chacune, et montrer que ce sont égaux.)

Maintenant que vous avez fini (!) vous avez peut-être constaté que c'est similaire à montrer un axiome pour un espace vectoriel. De plus, c'est très similaire au test de sous-espace.

Exemple: Trouver la matrice standard du S précédent.

On calcul alors T sur la base standard de $\mathbb{R}^4$:

$$S\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1+0 \\ 0-0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad S\left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0+1 \\ 0-0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$S\left(\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0+0 \\ 1-0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad S\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0+0 \\ 0-1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Alors la matrice standard de S est $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$. On a prouvé que :

$$S\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Note qu'on met les colonnes de la matrice standard dans l'ordre de la base standard.

3. Espaces

On peut définir LE NOYAU D'UNE TRANSFORMATION. C'est $\ker(T) = \{\mathbf{x} \mid T(\mathbf{x}) = \mathbf{0}\}$. De plus, L'IMAGE D'UNE TRANSFORMATION est $\mathrm{im}(T) = \{T(\mathbf{x}) \mid \mathbf{x} \in \mathbb{R}^n\}$.

Théorème 3. *Soit $T : \mathbb{R}^n \to \mathbb{R}^m$ une transformation linéiare, et A sa matrice standard.*

Alors $\ker(T) = \ker(A)$ *et* $\mathrm{im}(T) = \mathrm{col}(A)$.

La démonstration est directe : on observe que les définitions sont pareilles. Note que on a définit $\mathrm{col}(A)$ comme étant l'espace engendré par les colonnes de A. C'est exactment la même chose que dire toute combinaison linéaire des colonnes, et donc tout produit $A\mathbf{x}$. $\square$

Exemple: Soit $S : \mathbb{R}^4 \to \mathbb{R}^2$ défini par $S\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_3 - x_4 \end{bmatrix}$. Trouver une base pour $\ker(S)$.

On connaît la matrice de S : $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$. Donc $\ker(S) = \ker(A)$. Sans donner les détails, une base pour le noyau de cette matrice est $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Exercice: Soit $T : \mathbb{R}^2 \to \mathbb{R}^3$ défini par $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \\ x_2 - x_1 \end{bmatrix}$. Trouver une base pour $\mathrm{col}(T)$.

On connaît la matrice de T : $B = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix}$. Donc $\mathrm{im}(T) = \ker(B)$. Sans donner les détails, une base pour l'image est $\left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$. Aussi, on connait que

$$\mathrm{im}(T) = \mathrm{col}(B) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mid x_1 + x_2 = 0 \right\}$$

1. Projections

Soit X un sous-espace de $\mathbb{R}^n$. Soit $P : \mathbb{R}^n \to X$ la transformation défini par $P(\mathbf{u}) = \mathrm{proj}_X(\mathbf{u})$. On laisse comme exercice de montrer que c'est linéaire.

Exercice: Montrer que P est une transformation linéaire. (indice : écrire $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$, où $\mathbf{u}_1 \in X$ et $\mathbf{u}_2 \in X^\perp$ et noter que $P(\mathbf{u}) = \mathbf{u}_1$)

On trouvera la matrice standard de P de deux manières complètement différentes.

On écrit premièrement une matrice A (qui n'est pas la matrice standard !) tel que ses colonnes forment une base pour X. Ensuite on résoud $A^TA\mathbf{x} = A^T\mathbf{u}$. Puisque les colonnes de A sont indépendants (?), la matrice A^TA est inversible (c'était un devoir !) et donc on a $\mathbf{x} = \left(A^TA\right)^{-1} A^T\mathbf{u}$. Sachant la solution (unique !) $\mathbf{x}$, on calcul $P(\mathbf{u}) = \mathrm{proj}_X(\mathbf{u}) = A\mathbf{x} = A\left(A^TA\right)^{-1} A^T\mathbf{u}$.

Donc on a $P(\mathbf{u}) = M\mathbf{u}$ pour $M = A\left(A^TA\right)^{-1} A^T$. C'est la matrice standard de cette transformation.

Il reste des choses à vérifier. Certainement $P(\mathbf{u}) = M\mathbf{u}$, mais ce n'est pas la définition d'une matrice standard. On n'a pas du tout utilisé la définition pour arriver à cette matrice. Est-ce que cette méthode est valide ?

Voici une deuxième manière. Supposons qu'on a une base orthonormale $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ pour l'espace X.

On veut calculer la projection de $\mathbf{u}$ sur X, $\mathrm{proj}_X(\mathbf{u})$. Ce serait donc

$$\mathrm{proj}_X(\mathbf{u}) = (\mathbf{v}_1 \cdot \mathbf{u})\mathbf{v}_1 + (\mathbf{v}_2 \cdot \mathbf{u})\mathbf{v}_2 + \cdots + (\mathbf{v}_m \cdot \mathbf{u})\mathbf{v}_m$$

C'est une combinaison linéaire des vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$:

$$\mathrm{proj}_X(\mathbf{u}) = \left[\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_m \end{array}\right] \begin{bmatrix} \mathbf{v}_1 \cdot \mathbf{u} \\ \mathbf{v}_2 \cdot \mathbf{u} \\ \vdots \\ \mathbf{v}_m \cdot \mathbf{u} \end{bmatrix} = \left[\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_m \end{array}\right] \left[\begin{array}{c} \mathbf{v}_1 \\ \hline \mathbf{v}_2 \\ \hline \vdots \\ \hline \mathbf{v}_m \end{array}\right] \mathbf{u}$$

Si on met ces vecteurs comme colonnes d'une matrice B, on a donc $\mathrm{proj}_X(\mathbf{u}) = M\mathbf{u}$, avec $M = BB^T$.

Exercice: Trouver un lien entre ces deux méthodes "complètement différentes". (indice : qu'arrive-t-il si la base dans la première méthode est orthonormale ?)

On n'a encore pas utilisé la définition d'une matrice standard. Esst-ce que cette méthode est valide ? Est-ce possible d'avoir deux matrices standards *différents* ?

Il reste des explications ! Ces deux méthodes sont valides. La matrice standard est unique, donc tout méthode (correcte !) donne la même réponse. On laisse la discussion pour des cours plus avancés.

Supposons qu'on peut trouver une base orthonormale pour $\mathbb{R}^n$ qui contient la base pour U. donc $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots \mathbf{v}_n$ est une base orthonormale pour $\mathbb{R}^n$ (vous avez vu un tel exemple dans la question 5 du devoir 8). On voit que $P(\mathbf{v}_i) = \mathbf{v}_i$ si $1 \leq i \leq m$ et $P(\mathbf{v}_j) = \mathbf{0}$ si $m+1 \leq j \leq n$. Ce sont des vecteurs propres ; de plus c'est une base pour $\mathbb{R}^n$ formée de vecteurs propres. Bref : P possède deux espaces propres : U et $U^\perp$, correspondant aux valeurs propres 1 et 0.

Exercice: Vérifier que la matrice M possède les mêmes vecteurs propres que la transformation linéaire P. C'est-à-dire, calculer $M\mathbf{v}_i$ pour $1 \leq i \leq m$, et $M\mathbf{v}_j$ pour $m+1 \leq j \leq n$.

On peut prouver que n'importe quelle transformation linéaire qui est diagonalisable et ne possède que les valeurs propres 1 et 0 est une projection.

2. Graphiques

Exemple: Soit $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformation graphique qui correspond à faire une expansion d'un facteur de deux dans la direction de l'axe des x et de faire une contraction de un demi dans la direction de l'axe des y.

On cherche la matrice standard.

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}2\\0\end{bmatrix} \qquad\qquad T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\\frac{1}{2}\end{bmatrix}$$

La matrice est donc $\begin{bmatrix}2 & 0\\0 & \frac{1}{2}\end{bmatrix}$.

Exemple: Soit $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformation graphique qui correspond à faire une expansion d'un facteur de deux dans la direction du vecteur $\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right]$ et de faire une contraction de un demi dans la direction du vecteur $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$.

On cherche la matrice standard. On pourrait calculer l'effet de T sur la base standard, mais c'est meilleure d'utiliser les vecteurs propres.

Il existe une matrice standard, A. On connaît deux vecteurs propres de A :

$$A\begin{bmatrix}1\\2\end{bmatrix} = 2\begin{bmatrix}1\\2\end{bmatrix} \qquad\qquad A\begin{bmatrix}1\\0\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\2\end{bmatrix}$$

Il faut que les deux multiplicités algébriques soient un, car le degré du polynôme caractéristique est $n = 2$. Donc les deux multiplicités géométriques sont aussi un, et les vecteurs ci-haut sont chacune une base pour leur espace propre. La matrice A est diagonalisable et on connaît déja une base pour $\mathbb{R}^2$ formée de vecteurs propres.

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}^{-1}$$

On a trouvé la diagonalisation de A sans trouver A ! Si on avait a faire cette opération graphique sur un image, il s'agit de prendre chaque point comme vecteur $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ et de la multiplier par A.

Exemple: Soit $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformation graphique qui correspond à faire une rotation de θ degrés dans la direction contraire au sens de l'horloge.

On cherche la matrice standard.

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

Donc la matrice standard est $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

Un vecteur propre serait $\mathbf{x}$ tel que $A\mathbf{x} = \lambda\mathbf{x}$. Mais un image graphique suggère qu'il n'y a aucun tel $\mathbf{x}$: c'est une rotation, donc la direction de chaque vecteur non-nul change. Donc il n'y a aucun vecteur propre ? Par contre on a un théorème qui dit qu'il y a toujours des valeurs propres, et pour chacune, un espace propre de dimension au moins un. Hmmmm ?

Calculons les valeurs propres de A :

$$\det\left(\lambda I - \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\right) = (\lambda - \cos\theta)^2 + \sin\theta^2 = \lambda^2 - 2\cos\theta\lambda + 1$$

Les valeurs propres sont alors (formule quadratique) $\cos\theta \pm i \sin\theta$. Complexes !

Il n'y a aucun vecteur propre *réel* mais il y a des vecteurs propres : ce sont des vecteurs complexes. Pour $E_{\cos\theta + i\sin\theta}$ on a la base $\left[\begin{smallmatrix} 1 \\ -i \end{smallmatrix}\right]$, et pour $E_{\cos\theta - i\sin\theta}$ on a la base $\left[\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right]$. On a trouvé ces vecteurs de la manière standard, sauf avec une matrice complexe. Note que c'est complexe, mais pas compliqué : en fait, puisque les valeurs propres viennent en paires conjugés, les vecteur propres aussi !

Exercice: Soit λ une valeur propre de A, avec vecteur propre correspondant $\mathbf{x}$. Montrer que $\overline{\lambda}$ est aussi valeur propre, avec vecteur propre $\overline{\mathbf{x}}$.

Donc on a trouvé la base pour $E_{\cos\theta + i\sin\theta}$, et ensuite on a pris la conjugué de la réponse pour obtenir $E_{\cos\theta - i\sin\theta}$. Si c'était des valeurs propres réelles, on aurait dû trouver chaque base individuellement.

En générale, on peut souvent écrire la matrice standard d'une transformation directement en produit matricielle, en déduisant les vecteurs propres. C'est une technique très puissante. On n'a pas le temps d'en développer plus dans ce cours, mais c'est pour donner un avant-goût.