Final Exam Review

1. Linear Equations and Matrices

1.1: Matrices

- $A = B$ iff A and B have the same size and all coresponding entries are equal
- sum $(A + B)$: add all corresponding entries
- difference $(A - B)$: subtract all corresponding entries
- zero matrix (0): all entries are zeros
- negative matrix $(-A)$: negate every entry of A
- scalar product (cA): multiply every entry in A by scalar c
  NOTE:
  1. $c(A + B) = cA + cB$
  2. $(c + d)A = cA + dA$
  3. $c(dA) = (cd)A$
  4. $1A = A$
- if $cA = 0$, either $c = 0$ or $A = 0$
- transposition (A^T): "flip" A over it's main diagonal
- matrix A is symmetric iff $A = A^T$ and A is necessarily square

1.2: Linear Equations

- linear equation: includes the variables $x_1, x_2, \cdots, x_n$ and coefficients $a_1, a_2, \cdots, a_n$
- $AX = B$, where B is the constant term
- matrix of variables: $[X = x_1, x_2, ...x_n]^T$
- a set of numbers $s_1, s_2, \cdots, s_n$ is the solution to an equation if $a_1s_1 + a_2s_2 + ... + a_ns_n = b$ when we substitute $x_1 = s_1, x_2 = s_2, \cdots, x_n = s_n$
- system (of linear equations): holds a finite number of equations
- solution to the system: has a solution to *every* equation
- inconsistent system: has no solutions
- consistent system: has one or more solutions
- parameters are used when there are infinitely many solutions
- coefficient matrix: has only the coefficients of the system $(a_1, a_2, \cdots a_n)$
- augmented matrix: has the coefficients and the solutions of the system
- eqivalent systems: have the same solution(s)
  Elementary Row Operations:
  1. interchange two rows

2. multiply a row by a non-zero
3. add a multiple of one row to another

- row-echelon form (REF):
  1. all zero rows are at the bottom
  2. the first non-zero in a non-zero row is a 1
  3. each leading 1 is to the right of the one above it
- reduced row-echelon form (RREF): same as REF except each leading 1 is the only non-zero in it's row
- ever matrix can be carried to REF with elementary row operations
- leading variables: correspond to leading 1s
- $rankA$: number of leading 1s in A
- unique solution: every variable is a leading variable
- infinite solutions: there is at least one non-leading variable

1.3: Homogeneous Systems

- homogeneous system: all constants are zeros
- trivial solution: $x_1 = 0, x_2 = 0, \cdots x_n = 0$
- nontrivial solution: at least one variable is non-zero
- basic solutions: produced by Gaussian algorithm
- no basic solutions: iff the trivial solution is the only solution
  homogeneous system with n variables and rank r:
  1. algorithm produces $n - r$ basic solutions
  2. every solution is a linear combination of the basic solutions

1.4: Matrix Multiplication

- the dot product of $R = [r_1, r_2, \cdots r_n]^T$ and $C = [c_1, c_2, \cdots, c_n]^T$ is the number $r_1c_1 + r_2c_2 + ... + r_nc_n$
- product of $A_{(m \times n)}$ matrix and $B_{(n \times p)}$ matrix is an $m \times p$ matrix whose (i, j)-entry is the dot product of row i of A and column j of B
- two matrices commute if $AB = BA$
- identity matrix (I_n): square matrix with a main diagonal of 1s and filled in with 0s
- associative laws: $A(BC) = (AB)C$
- distributive laws: $A(B \pm C) = AB \pm AC$ and $(B \pm C)A = BA \pm CA$
- $AX = 0$ is the associated homogeneous matrix

1.5: Matrix Inverses

- if A is a square matrix, C is it's inverse if $AC = I$ and $CA = I$
- invertible matrix: any square matrix with an inverse
- if $AC = I = CA$ then $C = A^{-1}$
- determinant: $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $ad - bc$
- adjoint: $\begin{pmatrix} a & -b \\ -c & d \end{pmatrix}$
- from $AX = B$, $X = A^{-1}B$

- so, if the $n \times n$ coefficient matrix is invertible, the system has the unique solution $X = A^{-1}B$
- to find A^{-1}, take $\left(\begin{smallmatrix} A & I \end{smallmatrix} \right)$ and row reduce to $\left(\begin{smallmatrix} I & A^{-1} \end{smallmatrix} \right)$
- if A and B are invertible, so is AB and $(AB)^{-1} = B^{-1}A^{-1}$
- A^k is invertible for $k \geq 1$ and $(A^k)^{-1} = (A^{-1})^k$
- A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$
- cA is invertible if A is invertible and $c \neq 0$, written as $(cA)^{-1} = \frac{1}{c}A^{-1}$
- upper triangular: all zeros below the main diagonal
- lower triangular: all zeros above the main diagonal

2. Determinants and Eigenvalues

2.1: Cofactor Expansions

- A has an inverse if $detA \neq 0$
- $det\left(\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & i \end{smallmatrix} \right) = aei + bfg + cdh - ceg - afh - bdi$
- idea is to define the determinant of a 3×3 in terms of a 2×2 and the determinant of a 4×4 in terms of a 3×3 and so on
- $det\left(\begin{smallmatrix} a & b & c \\ d & e & f \\ g & h & i \end{smallmatrix} \right) = aei + bfg + cdh - ceg - afh - bdi$
  $= a(ei - fh) - b(di - fg) + c(dh - eg)$
  $= adet\left(\begin{smallmatrix} e & f \\ h & i \end{smallmatrix} \right) - bdet\left(\begin{smallmatrix} d & f \\ g & i \end{smallmatrix} \right) + cdet\left(\begin{smallmatrix} d & e \\ g & h \end{smallmatrix} \right)$
- given an $n \times n$ matrix A, let A_{ij} denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row i and column j, then defining the (i, j)-cofactor $C_{ij}(A) = (-1)^{i+j}det(A_{ij})$ for each i and j
- the signs for position are represented below
  $$\begin{pmatrix} +1 & -1 & +1 & -1 & \cdots \\ -1 & +1 & -1 & +1 & \cdots \\ +1 & -1 & +1 & -1 & \cdots \\ -1 & +1 & -1 & +1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$
- the determinant of any $n \times n$ matrix $A = a_{11}C_{11}(A) + a_{12}C_{12}(A) + ... + a_{1n}C_{1n}(A)$
- if a square matrix A has a row or column of entirely zeros, then $detA = 0$
- row and column operations reduce computing time for determinants
  1. if B is obtained from A by interchanging two different rows (or columns), then $detB = -detA$
  2. if B is obtained from A by multiplying a row (or column) by a number k, then $detB = kdetA$
  3. if B is obtained from A by adding a multiple of some row (or column) of A to a different row (or column), then $detB = detA$
- if a square matrix A has two identical rows or columns, then $detA = 0$
- if A is an $n \times n$ matrix, then $det(kA) = k^n detA$ for any scalar k
- if a square matrix A is triangular, then $detA$ is the product of the entries on the main diagonal
- the above implies that $detI = 1$ for any identity matrix I
- if A is any $n \times n$ matrix, then $det(A^T) = detA$

2.2: Determinants and Inverses

- if A and B are $n \times n$ matrices, then $det(AB) = detAdetB$
- $det(A^k) = (detA)^k$
- if A is invertible, then $det(A^{-1}) = \frac{1}{detA}$
- if A and B are square matrices (possibly of different sizes), then $det\left(\begin{smallmatrix} A & X \\ 0 & B \end{smallmatrix}\right) = detAdetB$ and $det\left(\begin{smallmatrix} A & 0 \\ X & B \end{smallmatrix}\right) = detAdetB$

2.3: Diagonalization and Eigenvalues

- diagonalize a matrix: find an invertible matrix P such that $P^{-1}AP = D$ is a diagonal matrix
- if $A = PDP^{-1}$, then $A^k = PD^kP^{-1}$ for each $k = 1, 2,$
- if A is an $n \times n$ matrix, a number λ is called an eigenvalue of A if $AX = \lambda X$ for some column $X = 0$
- such a nonzero colmn is called an eigenvector of A corresponding to the eigenvalue λ
- if I is the identity matrix of the same size as A, the homogeneous system $(\lambda I - A)X = 0$ has a nontrivial solution $X \neq 0$
- the above happens only if $(\lambda I - A)$ is not invertible, meaning that the determinant is 0 ($det(\lambda I - A) = 0$)
- if A is an $n \times n$ matrix, the characteristic polynomial $c_A(x)$ of A is defined by $c_A(x) = det(\lambda I - A)$
- the eigenvalues λ of A are the roots of the characteristic polynomial $c_A(x)$ of A
- the eigenvectors X corresponding to λ are the nonzero solutions to the homogeneous system $(\lambda I - A)X = 0$ of linear equations with $\lambda I - A$ as coefficient matrix
- the eigenvalues of a real matrix need not be real numbers
- an $n \times n$ matrix always has n (possibly complex) eigenvalues, but they may not be distinct
- the eigenvalues of a square matrix A usually are not computed as the roots of the characteristic polynomial
- diagonal matrix: all of the entries off of the main diagonal are zero, D has the form $D = \left(\begin{smallmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{smallmatrix}\right)$ where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are numbers
- if D and E are two diagonal matrices, their product DE and sum $D + E$ are again diagonal
- a square $n \times n$ matrix A is called diagonalizable if $P^{-1}AP$ is diagonal for some invertible $n \times n$ matrix P
- the invertible matrix P is called a diagonalizing matrix for A
- A is diagonalizable iff it has eigenvectors $X_1, X_2, \cdots, X_n$ such that the matrix $P = [X_1, X_2, \cdots, X_n]$ is invertible
- $P^{-1}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ where, for each i, λ_i is the eigenvalue of A corresponding to X_i
- if A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable
- an eigenvalue λ of a square matrix A is said to have multiplicity m if it occurs m times as a root of the characteristic polynomial $c_A(x)$
- a matrix is diagonalizable iff every eigenvalue λ of multiplicity m yields m basic eigenvectors
- the basic solutons of the system $(\lambda I - A)X =)$ become columns in the invertible diagonalizing matrix P such that $P^{-1}AP$ is diagonal
  to diagonalize an $n \times n$ matrix A:
  1. find the distinct eigenvalues λ of A
  2. compute basic eigenvectors corresponding to each of these eigenvalues λ from basic solutions of the homogeneous system $(\lambda I - A)X = 0$
  3. the matrix A is diagonalizable iff there are n basic eigenvectors in all
  4. if A is diagonalizable, the $n \times n$ matrix P with these basic eigenvectors as columns is a diagonalizing matrix for A; that is, P is invertible and $P^{-1}AP$ is diagonal

2.5: Complex Numbers

- $i^2 = -1$
- $a + ib$ where a and b are real
- a is called the real part and b is called the imaginary part
- if $a + bi = a' + b'i$, then $a = a'$ and $b = b'$
- the x-axis is the real axis and the y-axis is the imaginary axis
- if $z = A + bi$ and $w = a' + b'i$, then $z + w = (a + a') + (b + b')i$ and $zw = (aa' - bb') + (ab' + a'b)i$
- conjugate: $\bar{z} = a - bi$
  properties of the conjugate
  1. $\overline{(z \pm w)} = \bar{z} \pm \bar{w}$
  2. $\overline{zw} = \bar{z}\bar{w}$
  3. $\bar{\bar{z}} = z$
  4. z is real iff $\bar{z} = z$
- modulus/absolute value: $|z| = \sqrt{a^2 + b^2}$ represents the distance from z to the origin 0
  properties of the modulus
  1. $|z| \leq 0$ for all complex numbers z
  2. $|z| = 0$ iff $z = 0$
  3. $|zw| = |z||w|$
- $|z|^2 = z\bar{z}$ for every complex number z
- $|zw|^2 = (zw)\ \overline{(zw)} = zw\ \bar{z}\bar{w} = (z\bar{z})(w\bar{w}) = |z|^2|w|^2 = (|z|\ |w|)^2$
- quadratic formula: $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
- discriminant: $d = b^2 - 4ac$
- if $d > 0$, there are two real roots; if $d = 0$, there is one (repeated) real root; if $d < 0$, there is no real root and x is irreducible
- fundamental theorem of algebra: every nonconstant polynomial with complex coefficients had complex roots
- every complex polynomial $f(x)$ of degree $n \geq 1$ has the form $f(x) = u \cdot (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ where complex numbers $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the roots of $f(x)$ and need not be all distinct, and $u \neq 0$ is the coefficient of x^n in $f(x)$
- every $n \times n$ complex matrix A has n complex eigenvalues (possibly with some repeated)

3. Vector Geometry

3.1: Geometric Vectors

- for the vector $\vec{AB}$, A is called the tail of the vector and B is called the tip of the vector
- standard position: tail is at the origin 0
- $\vec{v} = \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right) = [xyz]^T$, $\vec{v}$ is in matrix form, x, y and z are components of $\vec{v}$, and $\vec{v}$ is the position vector of point $P = P(x, y, z)$
- zero vector: $\vec{0} = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right)$
- negative vector: $\vec{v} = \left(\begin{smallmatrix} -x \\ -y \\ -z \end{smallmatrix}\right)$
- sum of $\vec{v}$ and $\vec{w}$: $\vec{v} + \vec{w} = \left(\begin{smallmatrix} x+x_1 \\ y+y_1 \\ z+z_1 \end{smallmatrix}\right)$
- scalar product: $a\vec{v} = \left(\begin{smallmatrix} ax \\ ay \\ az \end{smallmatrix}\right)$

- $\vec{v} = \vec{w}$ iff $x = x_1, y = y_1$, and $z = z_1$
- $||\vec{v}|| = \sqrt{x^2 + y^2 + z^2}$
- $\vec{v} = \vec{0}$ iff $||\vec{v}|| = 0$
- $||a\vec{v}|| = |a| \; ||\vec{v}||$ for any scalar a
- vector equality: two vectors are equal as vectors (intrinsic) iff that are equal as matrices
- $\vec{0}$ is the only vector with a length of 0
- if $\vec{v}$ is a vector, then $\vec{-v}$ is the vector with the same length as $\vec{v}$ but opposite direction
- the vector $\vec{v} + \vec{w}$ is the vector in between $\vec{v}$ and $\vec{w}$ (parallelogram law) or the vector from the tip of $\vec{v}$ to the tail of $\vec{w}$ (tip-to-tail method)
- the vector $\vec{v} - \vec{w}$ is the vector from the tip of $\vec{w}$ to the tip of $\vec{v}$
- if P_1 and P_2 are two points, $\vec{P_1P_2} = \left(\begin{smallmatrix} x_2-x_1 \\ y_2-y_1 \\ z_2-z_1 \end{smallmatrix}\right)$
- the distance between P_1 and P_2 is $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$
- if $a\vec{v} \neq 0$, the direction of $a\vec{v}$ is the same as $\vec{v}$ if $a > 0$ and opposite to $\vec{v}$ if $a < 0$
- two nonzero vectors are called parallel if they have the same or opposite direction
- if $\vec{v}$ and $\vec{w}$ are parallel, each of $\vec{v}$ and $\vec{w}$ is a scalar multiple of the other and one of $\vec{v}$ and $\vec{w}$ is a scalar multiple of the other

3.2: Dot Product and Projections

- dot product, or scalar product of $\vec{v}$ and $\vec{w}$: $\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w} = x_1x_2 + y_1y_2 + z_1z_2$
- $\vec{v} \cdot \vec{w}$ is a number
- $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$
- $\vec{v} \cdot \vec{0} = 0 = \vec{0} \cdot \vec{v}$
- $\vec{v} \cdot \vec{v} = ||\vec{v}||^2$
- $(a\vec{v}) \cdot \vec{w} = a(\vec{v} \cdot \vec{w}) = \vec{v} \cdot (a\vec{w})$
- $\vec{v} \cdot (\vec{u} \pm \vec{w}) = \vec{v} \cdot \vec{u} \pm \vec{v} \cdot \vec{w}$
- is either $\vec{v}$ or $\vec{w}$ is the zero vector, the angle between them is undefined
- $h^2 = a^2 + b^2 - 2abcos\theta$ where h is the angle opposite θ
- if θ is the angle between the nonzero vectors $\vec{v}$ and $\vec{w}$, then $\vec{v} \cdot \vec{w} = ||\vec{v}|| \; ||\vec{w}|| \; cos\theta$
- therefore the formula for the cosine of the angle between $\vec{v}$ and $\vec{w}$ is $cos\theta = \frac{\vec{v} \cdot \vec{w}}{||\vec{v}|| \; ||\vec{w}||}$
- $\vec{v} \cdot \vec{w} > 0$ iff θ is acute
- $\vec{v} \cdot \vec{w} = 0$ iff θ is a right angle , meaning $\vec{v}$ and $\vec{w}$ are orthogonal
- $\vec{v} \cdot \vec{w} < 0$ iff θ is obtuse
- projection theorem: let $\vec{v}$ and $\vec{d} \neq 0$ be vectors
  1. the projection of $\vec{v}$ on $\vec{d}$ is given by: $proj_{\vec{d}}(\vec{v}) = (\frac{\vec{v} \cdot \vec{d}}{||\vec{d}||^2})\vec{d}$
  2. the vector $\vec{v} - proj_{\vec{d}}(\vec{v})$ is orthogonal to $\vec{d}$
  3. the vector $\vec{v}$ can be written uniquely in the form $\vec{v} = \vec{v_1} + \vec{v_2}$ where $\vec{v_1}$ is parallel to $\vec{d}$ and $\vec{v_2}$ is orthogonal to $\vec{d}$; in fact $\vec{v_1} = proj_{\vec{d}}(\vec{v})$ and $\vec{v_2} = \vec{v} - \vec{v_1}$

3.3: Lines and Planes

- vector equation of a line:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

- scalar equations of a line

$$\begin{aligned} x &= x_0 + ta \\ y &= y_0 + tb \\ z &= z_0 + tc \end{aligned}$$

- a nonzero vector $\vec{n}$ is called a normal to a plane if it is orthogonal to every vector in the plane
- vector equation of a plane:

$$\vec{n} \cdot (\vec{p} - \vec{p_0}) = 0$$

- scalar equation of a plane:

$$\begin{aligned} a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \\ ax + by + cz = ax_0 + by_0 + cz_0 \end{aligned}$$

- let $\vec{n} = [a\ b\ c]^T \neq \vec{0}$ be a fixed nonzero vector
  1. every plane with normal $\vec{n}$ has the equation $ax + by + cz = k$ for some constant k
  2. for every constant k, the graph of the equation $ax + by + cz = k$ is a plane with normal $\vec{n}$
- cross product or vector product

$$\vec{v} \times \vec{w} = \begin{bmatrix} y_1 z_2 - y_2 z_1 \\ -(x_1 z_2 - x_2 z_1) \\ x_1 y_2 - x_2 y_1 \end{bmatrix} = \begin{bmatrix} det\begin{pmatrix} y_1 & y_2 \\ z_1 & z_2 \end{pmatrix} \\ -det\begin{pmatrix} x_1 & x_2 \\ z_1 & z_2 \end{pmatrix} \\ det\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \end{bmatrix}$$

- if $\vec{v}$ and $\vec{w}$ are vectors in space, then $\vec{v} \times \vec{w}$ is a vector orthogonal to both $\vec{v}$ and $\vec{w}$ and $\vec{v} \times \vec{w} = 0$ iff $\vec{v}$ and $\vec{w}$ are parallel

3.5: The Cross Product

- $\vec{v} \times \vec{w} = det\begin{bmatrix} \vec{i} & x_1 & x_2 \\ \vec{j} & y_1 & y_2 \\ \vec{k} & z_1 & z_2 \end{bmatrix} = det\begin{pmatrix} y_1 & y_2 \\ z_1 & z_2 \end{pmatrix}\vec{i} - det\begin{pmatrix} x_1 & x_2 \\ z_1 & z_2 \end{pmatrix}\vec{j} + det\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}\vec{k}$
- if $\vec{u}$, $\vec{v}$ and $\vec{w}$ are vectors, then $\vec{u} \cdot (\vec{v} \times \vec{w}) = det[\vec{u}\ \vec{v}\ \vec{w}]$
- $\vec{u} \cdot (\vec{v} \times \vec{w})$ is called the scalar triple product
- for any nonzero vectors $\vec{v}$ and $\vec{w}$, the cross product $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$
- let $\vec{u}$, $\vec{v}$ and $\vec{w}$ be vectors
  1. $\vec{v} \times \vec{w}$ is a vector
  2. $\vec{v} \times \vec{0} = \vec{0} = \vec{0} \times \vec{v}$
  3. $\vec{v} \times \vec{v} = S\vec{0}$
  4. $\vec{w} \times \vec{v} = -(\vec{v} \times \vec{w})$
  5. $(a\vec{v}) \times \vec{w} = a(\vec{v} \times \vec{w}) = \vec{v} \times (a\vec{w})$ for any scalar a
  6. $\vec{v} \times (\vec{u} + \vec{w}) = (\vec{v} \times \vec{u}) + (\vec{v} \times \vec{w})$
- the Lagrange identity: if $\vec{v}$ and $\vec{w}$ are any two vectors, then $||\vec{v} \times \vec{w}|||^2 = ||\vec{v}||^2||\vec{w}||^2 - (\vec{v} \cdot \vec{w})^2$
- if three vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ are given, they determine a geometric solid called a parallelepiped
- the volume of the parallelepiped is hA where h is the length of the projection of $\vec{u}$ on $\vec{v} \times \vec{w}$

$$h = |\tfrac{\vec{u} \cdot (\vec{v} \times \vec{w})}{||\vec{v} \times \vec{w}||^2}|\ ||\vec{v} \times \vec{w}|| = \tfrac{|\vec{u} \cdot (\vec{v} \times \vec{w})|}{||\vec{v} \times \vec{w}||} = \tfrac{|\vec{u} \cdot (\vec{v} \times \vec{w})|}{A}$$

- the volume of the parallelepiped determined by three vectors $\vec{u}$, $\vec{v}$ and $\vec{w}$ is $|\vec{u} \cdot (\vec{v} \times \vec{w})| = |det[\vec{u}\ \vec{v}\ \vec{w}]|$
- right-hand rule: let $\vec{v}$ and $\vec{w}$ be nonzero vectors which are not parallel (so $\vec{v} \times \vec{w} \neq 0$), and let θ be the angle between $\vec{v}$ and $\vec{w}$. if the vector $\vec{v} \times \vec{w}$ is grasped in the right hand and the fingers curl around from $\vec{v}$ to $\vec{w}$ through the angle θ, the thumb points in the direction of $\vec{v} \times \vec{w}$

4. The Vector Space $\mathbb{R}^n$

4.1: Subspaces and Spanning

- $\mathbb{R}^n$ is the set of all n-vectors
- a set U of vectors in $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if it has the following three properties:
  1. the zero vector 0 is in U
  2. if X and Y are in U then $X + Y$ is also in U (closed under addition)
  3. if X is in U, then rX is also in U for any scalar r (closed under multiplication)
- the null space of A: $nullA = \{X in \mathbb{R}^n | AX = 0\}$
- the image of A: $imA = \{Y in \mathbb{R}^m | Y = AX for some X in \mathbb{R}^n\}$
- $nullA$ is the set of all solutions to the homogeneous system $AX + 0$, and imA is the set of all columns Y in $\mathbb{R}^m$ such that the system $AX = Y$ has a solution
- if $X_1, X_2, , X_k$ are any vectors in $\mathbb{R}^n$, then $span\{X_1, X_2, , X_k\}$ is a subspace of $\mathbb{R}^n$ which contains each of the vectors $X_1, X_2, , X_k$ and if the vectors $X_1, X_2, , X_k$ all lie in some subspace V, then $span\{X_1, X_2, , X_k\} \subseteq V$

4.2: Independent Sets of Vectors

- a set $\{X_1, X_2, , X_k\}$ of vectors is called linearly independent (or simply independent) if it satisfies the following condition: if $t_1X_1 + t_2X_2 + \; + t_kX_k = 0$, then $t_1 = t_2 = \; = t_k = 0$
- if $\{X_1, X_2, , X_k\}$ is an independent set of vectors, then every vector X in $span\{X_1, X_2, , X_k\}$ has a unique representation as a linear combination of the X^i
- a linear combination of vectors is trivial if every coefficient is zero
- if an $n \times n$ matrix A is invertible, then the columns of A are linearly independent in $\mathbb{R}^n$ and they span $\mathbb{R}^n$; the rows of A are linearly independent in $\mathbb{R}^n$ and they span $\mathbb{R}^n$
- a set $\{X_1, X_2, , X_k\}$ of vectors in $\mathbb{R}^n$ is linearly dependent (or simply dependent) if it is not linearly independent, ie iff one of the vectors X_i is a linear combination of the others
- if $\vec{v}$ and $\vec{w}$ are nonzero vectors in $\mathbb{R}^3$, then $\{\vec{v}, \vec{w}\}$ is linearly dependent iff $\vec{v}$ and $\vec{w}$ are parallel and $\{\vec{v}, \vec{w}\}$ is linearly independent iff $\vec{v}$ and $\vec{w}$ are not parallel

4.3: Fundamental Theorem

- if U is a subspace of $\mathbb{R}^n$ and U is spanned by m vectors and contains k linearly independent vectors, then $k \leq m$
- no linearly independent set in $\mathbb{R}^n$ can contain more than n vectors
- no spanning set for $\mathbb{R}^n$ can contain fewer than n vectors
- if U is a subspace of $\mathbb{R}^n$, a set $\{X_1, X_2, , X_k\}$ of vectors in U is called a basis of U if $\{X_1, X_2, , X_k\}$ is linearly independent and $U = span\{X_1, X_2, , X_k\}$
- invariance theorem: if $\{X_1, X_2, , X_k\}$ and $\{Y_1, Y_2, , Y_m\}$ are two bases of a subspace U in $\mathbb{R}^n$, then $k = m$
- suppose $\{X_1, X_2, , X_k\}$ is a linearly independent set of vectors in $\mathbb{R}^n$; if Y is any vector of $\mathbb{R}^n$ which is not in $span\{X_1, X_2, , X_k\}$, then the larger set $\{Y, X_1, X_2, , X_k\}$ is also linearly independent
- if $U \neq \{0\}$ is any nonzero subspace of $\mathbb{R}^n$, then U has a basis, and $dimU \leq n$, any linearly independent subset of U can be enlarged to a basis of U, and any spanning set of U contains a basis of U
- let U and V denote subspaces of $\mathbb{R}^n$
  1. if $U \subseteq V$, then $dimU \leq dimV$
  2. if $U \subseteq V$ and $dimU = dimV$, then $U = V$
  3. if $dimU = d$, then any set of d linearly independent vectors in U is automatically a basis of U
  4. if $dimU = d$, then any spanning set for U containing d vectors is automatically a basis of U

4.4: Rank

- the column space, $colA$, of an $m \times n$ matrix A is the subspace of $\mathbb{R}^m$ spanned by the columns of A
- the row space, $rowA$, of an $m \times n$ matrix A is the subspace of $\mathbb{R}^n$ spanned by the rows of A
- if $A \to B$ using a sequence of row operations, then $rowB = rowA$
- if $A \to B$ using a sequence of column operations, then $colB = colA$
- if A is an $m \times n$ matrix, then $rankA = dim(rowA)$ and if $A \to R$ by row operations where R is a row-echelon matrix, the nonzero rows of R are a basis of $rowA$
- rank theorem: if A is an $m \times n$ matrix, then $dim(rowA) = dim(colA) = rankA$ and if $A \to R$ by row operations where R is a row-echelon matrix, and if the leading 1s are in columns $j_1, j_2, \cdots, j_r$ of R, then the corresponding columns $j_1, j_2, \cdots, j_r$ of A are a basis of $colA$
- if A is an $m \times n$ matrix, then $rankA \leq m$ and $rankA \leq n$
- an $n \times n$ matrix A is invertible iff $rankA = n$
- $rankA^T = rankA$ for any matrix A
- if A is an $m \times n$ matrix, then $rankA = rank(UAV)$ for any invertible matrices U and V
- if A is any $m \times n$ matrix and $rankA = n$, then $AX = 0$, with X as a column, has only the trivial solution $X = 0$, the columns of A are linearly independent, and A^TA is an invertible $n \times n$ matrix
- if A is any $m \times n$ matrix and $rankA = m$, then $AX = B$ has a solution for every column B in $\mathbb{R}^m$, the columns of A span $\mathbb{R}^m$, and AA^T is an invertible $m \times m$ matrix
- if A is any $m \times n$ matrix of rank r, then $colA = imA = \{AX | Xin\mathbb{R}^n\}$ hence the rank theorem provides a basis for imA and in particular, $dim(imA) = rankA$
- if A is any $m \times n$ matrix, then $dimA(imA) + dim(nullA) = n$

4.5: Orthogonality

- if X, Y, and Z denote columns in $\mathbb{R}^n$, then:
  1. $X \cdot Y = Y \cdot X$
  2. $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$
  3. $(aX) \cdot Y = a(X \cdot Y) = X \cdot (aY)$ for all scalars a
  4. $||X||^2 = X \cdot X$
  5. $||X|| \geq 0$ and $||X|| = 0$ iff $X = 0$
  6. $||aX|| = |a| \; ||X||$ for all scalars a
- a vector E in $\mathbb{R}^n$ is called a unit vector if $||E|| = 0$
- cauchy inequality: if X and Y are vectors in $\mathbb{R}^n$, then $|X \cdot Y| \leq ||X|| \; ||Y||$ but the equality only holds if one of X and Y is a scalar multiple of the other
- triangle inequality: if X and Y are vectors in $\mathbb{R}^n$, then $||X + Y|| \leq ||X|| + ||Y||$
- if X, Y, and Z denote vectors in $\mathbb{R}^n$, then $d(X,Y) \geq 0$ for all X and Y, $d(X,Y) = 0$ only if $X = Y$, $d(X,Y) = d(Y,X)$, and from the triangle inequality $d(X,Y) \leq d(X,Z) + d(Z,Y)$
- pythagoras' theorem: if $\{X_1, X_2, \cdots, X_k\}$ is an orthogonal set of vectors in $\mathbb{R}^n$, then $||X_1 + X_2 + \; + X_k||^2 = ||X_1||^2 + ||X_2||^2 + ... + ||X_k||^n$
- every orthogonal set of vectors in $\mathbb{R}^n$ is linearly dependent
- expansion theorem: if $\{X_1, X_2, \cdots, X_k\}$ is an orthogonal basis of a subspace U of $\mathbb{R}^n$ and if X is any vector in U then, $X = \frac{X \cdot X_1}{||X_1||^2} X_1 + \frac{X \cdot X_2}{||X_2||^2} X_2 + ... + \frac{X \cdot X_k}{||X_k||^2} X_k$

- orthogonal lemma: if $\{F_1, F_2, \cdots, F_k\}$ is an orthogonal set in $\mathbb{R}^n$ and given X in $\mathbb{R}^n$ write $F_{k+1} = X - (\frac{X \cdot F_1}{||F_1||^2}F_1 + \frac{X \cdot F_2}{||F_2||^2}F_2 + ... + \frac{X \cdot F_k}{||F_k||^2}F_k)$ then F_{k+1} is orthogonal to each of $F_1, F_2, \cdots, F_k$ and if X is not in $span\{F_1, F_2, \cdots, F_k\}$, then $F_{k+1} \neq 0$ and the larger set $\{F_1, F_2, \cdots, F_k, F_{k+1}\}$ is an orthogonal set

- if U is a nonzero subspace of $\mathbb{R}^n$, then every orthogonal subset $\{F_1, F_2, \cdots, F_k\}$ of U is part of an orthogonal basis of U and in particular, U had an orthogonal basis

- gram-schmidt algorithm: if $\{X_1, X_2, \cdots, X_m\}$ is any basis of a subspace U of $\mathbb{R}^n$, construct $F_1, F_2, \cdots, F_m$ in U successively as follows

$$
\begin{gathered}
F_1 = X_1 \\
F_2 = X_2 - \tfrac{X \cdot F_1}{||F_1||^2}F_1 \\
F_3 = X_3 - \tfrac{X \cdot F_1}{||F_1||^2}F_1 - \tfrac{X \cdot F_2}{||F_2||^2}F_2 \\
\vdots \\
F_k = X_k - \tfrac{X \cdot F_1}{||F_1||^2}F_1 - \tfrac{X \cdot F_2}{||F_2||^2}F_2 - \cdots - \tfrac{X \cdot F_{k-1}}{||F_{k-1}||^2}F_{k-1}) \\
\vdots
\end{gathered}
$$

  for each $k = 2, 3, \cdots, m$. Then:

  1. $F_1, F_2, \cdots, F_m$ is an orthogonal basis of U
  2. $span\{F_1, F_2, \cdots, F_k\} = span\{X_1, X_2, \cdots, X_k\}$ for each $k = 1, 2, \cdots, m$

4.6: Projections and Approximation

- orthogonal complement: $U^\perp$ is defined to be the set of all vectors that are orthogonal to every vector in U: $U^\perp = \{X$ in $\mathbb{R}^n | X \cdot Y = 0$ for all Y in $U\}$

- if U denotes a subspace of $\mathbb{R}^n$, then $U^\perp$ is also a subspace of $\mathbb{R}^n$ and if $U = span\{X_1, \cdots, X_k\}$, then $U^\perp = \{X | X \cdot X_i = 0$ for each $i = 1, 2, \cdot, k\}$

- if U is a nonzero subspace of $\mathbb{R}^n$, and X is any vector in $\mathbb{R}^n$, and if $\{F_1, F_2, \cdots, F_m\}$ is an orthogonal basis of U, then the vector $P = \frac{X \cdot F_1}{||F_1||^2}F_1 + \frac{X \cdot F_2}{||F_2||^2}F_2 + ... + \frac{X \cdot F_m}{||F_m||^2}F_m$ depends only on U and X; that is, given X, P is the same for every choice of orthogonal basis $\{F_1, F_2, \cdots, F_m\}$

- projection theorem: if U is a subspace of $\mathbb{R}^n$ and X is a vector in $\mathbb{R}^n$, then $proj_U(X)$ is in U and $X - proj_U(X)$ is in $U^\perp$

- if U is a subspace of $\mathbb{R}^n$ and X is a vector in $\mathbb{R}^n$, then if $X = Y + Z$ where Y is in U and Z is in $U^\perp$, then necessarily $Y = proj_U(X)$ and $Z = X - proj_U(X)$

- if U is any subspace of $\mathbb{R}^n$, then $dim(U^\perp) = n - dim U$ and $U^{\perp\perp} = U$

- if U is any subspace of $\mathbb{R}^n$ and X is a vector in $\mathbb{R}^n$, then $X = proj_U(X) + proj_{U^\perp}(X)$

- approximation theorem: let U be a subspace of $\mathbb{R}^n$ and X be a vector in $\mathbb{R}^n$, and write $P = proj_U(X)$; then P is the vector in U closest to X in the sense that $||X - P|| < ||X - Y||$ for all $Y \neq P$ in U

4.9: Linear Transformations

- a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if it satifies the following two properties:

  1. $T(X + Y) = T(X) + T(Y)$ for all X and Y in $\mathbb{R}^n$ (T preserves addition)
  2. $T(aX) = aT(X)$ for all X in $\mathbb{R}^n$ and all scalars a (T preserves scalar multiplication)

- $\mathbb{R}^n \to \mathbb{R}^m$ is called a linear operator on $\mathbb{R}^n$

- let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, let $X, X_1, X_2, \cdots, X_k$ denote vectors in $\mathbb{R}^n$, and let $a_1, a_2, \cdots, a_k$ denote scalars; then $T(0) = 0$, $T(-X) = -T(X)$, $T(a_1X_1 + a_2X_2 + \cdots + a_kX_k) = a_1T(X_1) + a_2T(X_2) + \cdots + a_kT(X_k)$

- if A is an $m \times n$ matrix and $AX = 0$ for all X in $\mathbb{R}^n$, then $A = 0$

- let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, then T is the matrix transformation induced by a uniquely determined $m \times n$ matrix A: $T(X) = AX$ for all X in $\mathbb{R}^n$

- furthermore, A is given in terms of its columns by $A = [T(E_1)\ T(E_2)\ \cdots\ T(E_n)]$ where $\{E_1, E_2, \cdots, E_n\}$ is the standard basis of $\mathbb{R}^n$.
- the matrix A above is called the standard matrix of T
- let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^k$ be linear transformations, then the composite $S \circ T : \mathbb{R}^n \to \mathbb{R}^k$ is also a linear transformation, and moreover, if S and T have standard matrices A and B respectively, then $S \circ T$ has standard matrix AB
- let $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ be two linear transformations and suppose that $\mathbb{R}^n = span\{Y_1, Y_2, \cdots, Y_k\}$, and if $S(Y_i) = T(Y_i)$ for each i, then $S = T$
- let $\mathcal{F} = \{F_1, F_2, \cdots, F_n\}$ be any basis of $\mathbb{R}^n$, and select any n vectors $\{Z_1, Z_2, \cdots, Z_n\}$ in $\mathbb{R}^m$, possibly not distinct; then there is a unique linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that $T(F_i) = Z_i$ holds for each $i = 1, 2, \cdots, n$
- in fact, T can be defined as follows: if X is any vector in $\mathbb{R}^n$, express X as a linear combination of the vectors in the basis $\mathcal{F}$, say $X = a_1F_1 + a_2F_2 + ... + a_nF_n$; then $T(X)$ is given by $T(X) = a_1Z_1 + a_2Z_2 + ... + a_nZ_n$
- if S and S_1 are both inverses of T, then $S = S_1$
- let $T : \mathbb{R}^n \to \mathbb{R}^m$ be an operator with standard matrix A; then T has an inverse iff A is an invertible matrix, and in this case A^{-1} is the standard matrix of T^{-1}
- let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator, and let $\mathcal{F}$ be a basis of $\mathbb{R}^n$; then a uniquely determined $n \times n$ matrix $M_{\mathcal{F}}(T)$ exists such that $C_{\mathcal{F}}[T(X)] = M_{\mathcal{F}}(T)C_{\mathcal{F}}(X)$ for all X in $\mathbb{R}^n$
- furthermore, if $\mathcal{F} = \{F_1, F_2, \cdots, F_n\}$ and we put $P = [F_1\ F_2\ \cdots\ F_n]$, we have
  1. $M_{\mathcal{F}}(T) = P^{-1}AP$ where A is the standard matrix of P
  2. $M_{\mathcal{F}}(T) = [C_{\mathcal{F}}(T(F_1))\ C_{\mathcal{F}}(T(F_2))\ \cdots\ C_{\mathcal{F}}(T(F_n))]$
  3. $M_{\mathcal{F}}(T)$ is invertible iff T is invertible

5. Vector Spaces

5.1: Examples and Basic Properties

- axioms for vector addition:
  1. if v and w are in V, then $v + w$ is in V
  2. $v + w = w + v$ for all v and w in V
  3. $u + (v + w) = (u + v) + w$ for all u, v and w in V
  4. an element 0 in V exists such that $v + 0 = v$ for all v in V
  5. for each v in V an element $-v$ exists such that $v + (-v) = 0$
- axioms for scalar multiplication:
  1. if v is in V, then av is in V for all scalars a
  2. $a(v + w) = av + aw$ for all v and w in V and all scalars a
  3. $(a + b)v = av + bv$ for all v in V and all scalars a and b
  4. $a(bv) = (ab)v$ for all v in V and all scalars a and b
  5. $1v = v$ for all v in V
- if $p(x)$ is a nonzero polynomial, the highest power of x in $p(x)$ with nonzero coefficient is called the degree of $p(x)$, and is denoted as $deg\ p(x)$; and the coefficient of this highest power is called the leading coefficient of $p(x)$
- cancellation: if $w + v = u + v$ in a vector space, then $w = u$
- let v denote a vector in a vector space V, and let a denote a scalar:
  1. $0v = 0$
  2. $a0 = 0$
  3. if $av = 0$, then either $a = 0$ or $v = 0$
  4. $(-1)v = -v$

5. $(-a)v = -(av) = a(-v)$

- subspace test: a subset U of a vector space V is a subspace iff it satisfies the following three conditions:
  1. 0 is in U
  2. if u is in U, then au is in U for all scalars a (U is closed under scalar multiplication)
  3. if u and u_1 are in U, then $u + u_1$ is in U (U is closed under addition)
- let $U = span\{v_1, v_2, \cdots, v_n\}$ in a vector space V, U is a subspace of V containing each of the vectors $v_1, v_2, \cdots, v_n$ and U is the "smallest" subspace containing these vectors in the sense that any subspace of V that contains each of $v_1, v_2, \cdots, v_n$ must already contain U

5.2: Independence and Dimension

- to check if a set $\{v_1, v_2, \cdots, v_n\}$ is independent set a linear combination equal to zero: $r_1v_1 + r_2v_2 + \cdots + r_nv_n = 0$; and show somehow that this implies that all the coefficients are zero: $r_i = 0$ for each i
- fundamental theorem: suppose that V is a vector space and $V = span\{v_1, v_2, \cdots, v_n\}$, if $\{u_1, u_2, \cdots, u_m\}$ is any linearly independent subset of V, then $m \leq n$
- invariance theorem: if $\{b_1, b_2, \cdots, b_n\}$ and $\{d_1, d_2, \cdots, d_n\}$ are two bases of a vector space V, then $n = m$
- independent lemma: suppose that $\{v_1, v_2, \cdots, v_n\}$ is a linearly independent set of vectors in a vector space V, if v is any vector in V that is not in $span\{v_1, v_2, \cdots, v_n\}$, then $\{v, v_1, v_2, \cdots, v_n\}$ is also linearly independent
- dependent lemma: a set $\{v_1, v_2, \cdots, v_n\}$ of vectors in a vector space V is linearly dependent iff one of the v_i is a linear combination of the rest
- if $V \neq 0$ is a finite dimensional vector space amf V is spanned by n vectors, then V has a basis and $dimV \leq n$, every linear independent subset of V is part of a basis, and every finite spanning set of V contains a basis
- if V is a vector space with $dimV = n$ and $\mathcal{B}$ is a set of exactly n vectors of V, then $\mathcal{B}$ is independent iff $\mathcal{B}$ spans V; and in either case, $\mathcal{B}$ is a basis of V
- if U is a subspace of a vector space V and $dimV = n$, then U has a basis and $dimU \leq n$, if $dimU = n$, then $U = V$, and every basis of U is a part of a basis of V

5.3: Linear Transformations

- a transformation $T : V \to W$ is called a linear transformation if it preserves addition and scalar multiplication according to the following axioms:
  1. $T(v + v_1) = T(v) + T(v_1)$ for all vectors v and v_1 in V
  2. $T(rv) = rT(v)$ for all vectors v and all scalars r
- if $T : V \to W$ is a linear transformation, then $T(0) = 0$, $T(-v) = -T(v)$ for all v in V, and $T(r_1v_1+r_2v_2+\cdots+r_kv_k) = r_1T(v_1) + r_2T(v_2) + \cdots + r_kT(v_k)$ for all r_1 in $\mathbb{R}$ and all v in V
- let $S : V \to W$ and $T : V \to W$ be two linear transformations, and assume that $V = span\{v_1, v_2, \cdots, v_k\}$; if $S(v_i) = T(v_i)$ for each i, then $S = T$
- if $\{b_1, b_2, \cdots, b_n\}$ is a basis of a vector space V, every vector v in V can be uniquely represented as a linear combination $v = r_1b_1 + r_2b_2 + \cdots + r_nb_n$ where the r_i are in $\mathbb{R}$
- let $\{b_1, b_2, \cdots, b_n\}$ be a basis of a vector space V, and let $\{w_1, w_2, \cdots, w_n\}$ be arbitrary vectors in W (possibly not distinct); then there exists a unique linear transformation $T : V \to W$ such that $T(b_i) = w_i$ for each $i = 1, 2, \cdots, n$
- furthermore, the action of T is given as follows: if $v = r_1b_1 + r_2b_2 + \cdots + r_nb_n$ is in V where each r_i is in $\mathbb{R}$, then $T(v) = r_1w_1 + r_2w_2 + \cdots + r_nw_n$
- let $T : V \to W$ be a linear transformation; then $kerT$ is a subspace of V and imT is a subspace of W
- if $T : V \to W$ is a linear transformation, we say that T is onto if $imT = W$ and T is one-to-one if $T(v_1) = T(v_2)$ implies that $v_1 = v_2$ i.e. iff $kerT = \{0\}$
- dimension theorem: let $T : V \to W$ be a linear transformation; if both $kerT$ and imT are finite dimensional, then V is also finite dimensional and $dimV = dim(imT) + dim(kerT)$
- let $T : V \to W$ be a linear transformation; if $dimV = dimW$ is finite, then T is one-to-one iff T is onto