Solutions

- *No Solution*: The system of equations is *inconsistent* (i.e. contains a row/equation of the form [0 0 … 0 | *])
- *Infinite Solutions*: The system doesn't contain enough information (equations) to have a unique solution, resulting in *free variables*.
- *Unique Solution*: There is enough information (i.e. there is an equal or larger number of equations than variables; $n \geq m$) to determine the unique solution to a particular system, i.e. **NO** free variables.

Rank

- rank(A) = rk(A) = # of leading 1s in rref(A)
- Also represents the number of linearly independent columns of A, a.k.a. the *dimension* of the vector space generated by A's columns
- **Properties**
  - rk(A) ≤ n and rk(A) ≤ m => rk(A) ≤ min(n, m)
  - If A is inconsistent, then rk(A) < n since at least one row will consist of all zeros
  - **If A has a unique solution, then rank(A) = m, since every variable must be bounded to a value (non-free)**

Linear Transformations

- A function $T: \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation if $T(\vec{x}) = A\vec{x}$ for some matrix A of size $n \times m$
- T must also satisfy two properties:
  - $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$ & $T(k\vec{x}) = kT(\vec{x})$
- When given a linear transformation T, we can sometimes solve for A by multiplying it with the $m \times m$ Identity matrix to get:
  - A = $\begin{pmatrix} | & \cdots & | \\ T(e_1) & \ddots & T(e_m) \\ | & \cdots & | \end{pmatrix}$

Matrix Multiplication

* To multiply a $m \times n$ with a $r \times s$ matrix, *n* must equal *r* (*n* = *r*), and the resulting matrix will be $m \times s$

If $A = \begin{bmatrix} a11 & \cdots & a1m \\ \vdots & \ddots & \vdots \\ an1 & \cdots & anm \end{bmatrix}$ and $B = \begin{bmatrix} b11 & \cdots & b1s \\ \vdots & \ddots & \vdots \\ br1 & \cdots & brs \end{bmatrix}$,

then BA = $\begin{bmatrix} B\begin{matrix} a11 \\ \vdots \\ an1 \end{matrix} & \cdots & B\begin{matrix} a1m \\ \vdots \\ anm \end{matrix} \end{bmatrix}$

- The product of two matrices is simply the composition of their linear transformations
- **Properties**
  - Associativity
    - (AB)C=A(BC)
  - Distributivity
    - (A+B)C = AC+BC
    - A(C+D) = AC+AD

Dimension

- The number of vectors in a basis of V is called the dimension of V.
  - We can find at most *m* linearly independent vectors in V.
  - We need at least *m* vectors to span V.
  - If *m* vectors in V are linearly independent, then they form a basis of V.
  - If *m* vectors in V span V, then they form a basis of a.
- **Fundamental Theorem of Linear Algebra**
  - dim(im A) = rank(A)
  - dim(ker A) + dim(im A) = m for any A_{nxm}

Coordinates

- Consider a basis $\mathfrak{B} = (v_1, \dots, v_n)$ of subspace V of $\mathbb{R}^n$.

$$\vec{x} = c_1 v_1 + \cdots + c_n v_n$$

Thus, the scalars $c_1, c_2, \dots, c_n$ are the $\mathfrak{B}$-coordinates of x. In other words, $[\vec{x}]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ \cdots \\ c_n \end{bmatrix}$, meaning that $\vec{x} = S[\vec{x}]_{\mathfrak{B}}$, where S is $[c_1 \;\; \dots \;\; c_2]$

- Consider a linear transformation T. Let B be the $\mathfrak{B}$-matrix of T, and let A be the standard matrix of T. Then,

$$AS = SB \text{ or } B = S^{-1}AS$$

(**Similar matrices** are matrices that satisfy this property).

- A $\mathfrak{B}$-matrix of T is **diagonal** only if under its basis, T(v_1)= $c_1 v_1$, …, and T(v_n) = $c_n v_n$

Common 2D Linear Transformations

- Scaling
  - If a vector $\vec{v} \in \mathbb{R}^n$ is scaled by a factor of k, then the transformation matrix A is the $n \times n$ identity matrix × k
  - $\to k\begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots 1 & \vdots \\ 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} k & \cdots & 0 \\ \vdots & \ddots k & \vdots \\ 0 & \cdots & k \end{bmatrix}$
- Orthogonal Projection
  - If given a line L which is spanned by the vector $\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$, then $proj_L(\vec{x}) = x^{//} = k\vec{w} = \frac{\vec{x}\cdot\vec{w}}{||\vec{w}||^2}\vec{w}$
  - More generally, $proj_L(\vec{x}) = \frac{1}{w_1^2+w_2^2}\begin{bmatrix} w_1^2 & w_1w_2 \\ w_1w_2 & w_2^2 \end{bmatrix}\vec{x}$
- Reflection
  - If given a line L which is spanned by the vector $\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$, then $ref_L(\vec{x}) = \vec{x} - 2\vec{x}^{\perp} = 2\vec{x}^{//} - \vec{x} = \frac{1}{w_1^2+w_2^2}\begin{bmatrix} w_1^2 - w_2^2 & 2w_1w_{2s} \\ 2w_1w_2 & w_2^2 - w_1^2 \end{bmatrix}\vec{x}$
  - More generally, $ref_L(\vec{x}) = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}\vec{x}$
- Rotation (counter-clockwise)
  - $rot_\theta(\vec{x}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$
  - More generally,

$$rot_\theta(\vec{x}) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \; where\; a^2 + b^2 = 1$$

- Shearing
  - Vertical: Stretching a vector along the y-axis → $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$
  - Horizontal: Stretching a vector along the x-axis → $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
  - *Note: k does NOT represent the number of units shifted*

Common 3D Linear Transformations

- A reflection in the xy plane is given by:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ -z \end{bmatrix}$$

- Logic extends to other planes along axes.

Basis and Unique Representation

- A set of vectors form a basis of V iff every vector v in V can be expressed as a unique linear combination

Orthonormal Bases

- A basis in which all vectors are perpendicular and have a norm (length) of one.

Orthogonal Projections

- If a subspace V has an orthonormal basis $u_1, \dots, u_n$, then $proj_V(\vec{x}) = (u_1 \cdot x)u_1 + \cdots + (u_n \cdot x)u_n$

Orthogonal Complement

- $V^{\perp}$ of V is the set of those vectors in $\mathbb{R}^n$ that are orthogonal to all vectors in V.
- $V^{\perp} = \ker(proj_V(\vec{x}))$, $(V^{\perp})^{\perp} = V$, $V \cap V^{\perp} = \{\vec{0}\}$
- $\dim(V) + \dim(V^{\perp}) = n$

Cauchy-Schwarz Inequality

- $|x \cdot y| \leq ||x||\,||y||$

Gram-Schmidt Process and QR Factorization

- A method of constructing an orthonormal basis of some subspace
- M = $\begin{bmatrix} | & | \\ v_1 & v_2 \\ | & | \end{bmatrix} = QR = \begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix}\begin{bmatrix} ||v_1|| & u_1 \cdot v_2 \\ 0 & ||v_2|| \end{bmatrix}$
- **Pattern continues for larger number of vectors**

Orthogonal Transformations

- Orthogonal Transformations are linear transformations that preserve length.
- **Properties of Orthogonal Transformations**
  - ||Ax|| = ||x|| for all x in $\mathbb{R}^n$
  - Columns of A form an orthonormal basis of $\mathbb{R}^n$
  - $A^TA = I_n$.
  - $A^{-1} = A^T$.
  - $(Ax)\cdot(Ay) = x \cdot y$
- **Matrix of Orthogonal Projection**
  - Given a subspace V with an orthonormal basis, the matrix for the orthogonal projection onto V is

$$P = QQ^T, where\; Q = \begin{bmatrix} | & | & | \\ u_1 & \dots & u_n \\ | & | & | \end{bmatrix}$$

Inverses

- Set up matrix equation as:

$$\left[\begin{matrix} * & * & * \\ * & * & * \\ * & * & * \end{matrix}\middle|\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}\right]$$

and get the left-hand side in rref. The right-hand side will then represent the inverse matrix, assuming there is no inconsistency.

- **Invertibility Criterion**

  A is invertible if and only if …
  - $A\vec{x} = \vec{b}$ has a unique solution for any $\vec{b}$
  - $rank(A) = n$
  - $det(A) \neq 0$
  - $ker(A) = \{\vec{0}\}$
  - $im(A) = \mathbb{R}^n$
  - The column vectors of A form a basis of $\mathbb{R}^n$
  - The column vectors of A span $\mathbb{R}^n$
  - The column vectors of A are linearly independent.
  - 0 fails to be an eigenvalue of A.
- *Note that following from these properties, only square matrices may be invertible*

Subspaces

- $V \subset \mathbb{R}^n$, V is a subset of a vector space if and only if it satisfies:
  - $\vec{0}$ is in V
  - (sum closure) $\forall \vec{v}, \vec{w} \in V,\; \vec{v} + \vec{w} \in V$
  - (scalar closure) $\forall v \in V, \forall k \in \mathbb{R},\; k\vec{v} \in V$

Redundancy and Linear Independence

- A vector is *redundant* in a set of other vectors if it can be expressed as a linear combination of the other vectors, or it satisfies a *non-trivial linear relation*.
- A set of vectors is linearly independent if it does not contain any redundant vectors
- **Determining Linear Dependence**
  - Put the vectors into an augmented matrix by column and solve for its *rref*. If the last column contains a non-zero entry, then the set is <u>linearly dependent</u>
    - *This method corresponds to solving for the coefficients that satisfy the linear combination*
- **Zero Component Condition**
  - A vector is non-redundant if it contains a non-zero entry in a component where all the preceding vectors have a 0.
  - If this was the case for all vectors in a set, then the set is linearly independent

Properties of Determinant

- The determinant of a triangular matrix is just the product of the diagonal entries
- $\det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det\begin{pmatrix} A & 0 \\ B & C \end{pmatrix} = \det(A)\det(C)$
- $\det(A) = \det(A^T)$
- $\det(AB) = \det(A)\det(B)$
- $\det(A^m) = \det(A)^m$
- For similar matrices: $\det(A) = \det(B)$
- **Row Operations:**
  - Multiplying a single row by a scalar *k*:
    - $\det(B) = k * \det(A)$
  - Row Swap:
    - $\det(B) = -\det(A)$
  - Adding multiples of rows:
    - $\det(B) = \det(A)$

Least Squares and Data Fitting

- Normal Equation of Ax = b :
  - $A^TAx = A^Tb$
- Least-Squares Solution:
  - $x^* = (A^TA)^{-1}A^Tb$
    - Only unique in the case where A is linearly independent.
- Error:
  - $||b - Ax^*||$
- Orthogonal Projection onto V of basis $v_1, \dots, v_n$:
  - $A = \begin{bmatrix} | & | & | \\ v_1 & \dots & v_n \\ | & | & | \end{bmatrix}$
  - Then the matrix of the orthogonal projection onto V is $(A^TA)^{-1}A^T$
  - <u>The projection of *b* onto Im(*A*) is Ax^*</u>

Eigenvalues and Eigenvectors

- Given some matrix A, if a non-zero vector v satisfies $Av = \lambda v$, where λ is some scalar, then it is said that v is an *eigenvector* of A, and λ is the eigenvalue of that vector.
  - Geometrically, this is such a vector such that applying the transformation A keep the vector on the same line as the original vector.
- **Algebraic Method**
  - $Av = \lambda v \Rightarrow (A - \lambda I)v = 0 \Rightarrow$ The set of eigenvectors of A form the null space of $(A - \lambda I)v \Rightarrow \ker(A - \lambda I)$ is non-trivial, meaning $(A - \lambda I)$ is not invertible and $\det(A - \lambda I) = 0$
  - Thus, first solve for every value of λ using the determinant of A (the determinant will give the *characteristic polynomial*, whose roots will be the eigenvalues). Then apply each value to $(A - \lambda I)$ and solve for $\ker(A - \lambda I)$.
    - For a 2x2 matrix, the characteristic polynomial is given by:
    $$\lambda^2 - tr(A)\lambda - \det(A) = 0$$
    where *tr(A)* is given by sum of the diagonal elements of A.
- **Diagonalization**
  - The matrix A is diagonalizable iff there exists an eigenbasis for it. If $v_1, v_2, \ldots, v_n$ form an eigenbasis for A, such that $Av_1 = \lambda_1 v_1$, $Av_2 = \lambda_2 v_2$, …, $Av_n = \lambda_n v_n$, then the matrices:
  $$S = \begin{bmatrix} | & | & | \\ v_1 & v_2 & \ldots \\ | & | & | \end{bmatrix}, \quad B = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \vdots \end{bmatrix}$$
  will diagonalize A such that $S^{-1}AS = B$
    - Note: A might not have enough eigenvectors to form an eigenbasis, in which case A is NOT diagonalizable

Symmetric Matrices

- Symmetric Matrices are matrices that satisfy the property $A = A^T$.
- *Spectral Theorem*: A matrix A is *orthogonally diagonalizable* (there exists an orthonormal eigenbasis S for A such that $S^{-1}AS = S^TAS$ is diagonal) iff A is a symmetric matrix.
  - If two eigenvectors of A have *distinct* eigenvalues, then those two vectors must be orthogonal to each other.
- **Orthogonal Diagonalization of a Symmetric Matrix**
  - Use the Algebraic method to determine A's eigenvalues and find the basis of each eigenspace.
  - Use the Gram-Schmidt Process to find an orthonormal basis of each eigenspace.
  - Form an orthonormal eigenbasis $v_1, \ldots, v_n$ for A by concatenating the orthonormal bases found in step 2, and let
  $$S = \begin{bmatrix} | & | & | & | \\ v_1 & v_2 & \ldots & v_n \\ | & | & | & | \end{bmatrix}$$
  where S is orthogonal and $S^{-1}AS$ is diagonal.
- **Properties of Symmetric Matrices**
  - $A = A^T$
  - A^{-1} must also be symmetric

Miscellaneous Properties

- $\dim(Im(A)^{\perp}) = \dim(\ker(A^T))$
- $\ker(A^TA) = \ker(A)$

RREF

- A matrix is said to be in Reduced Row-Echelon Form if it satisfies the following:
  - The leading coefficient in each row is 1
  - The leading variables in each equation do not appear in others (i.e. rows with leading variables must have the rest of the column of the leading coefficient be 0)
  - Leading variables appear in increasing order (i.e. leading coefficient "move" right)

Dynamical Systems

- A dynamical system *x(t)* is a system such that:
$$x_1(t+1) = a_{1,1}x_1(t) + a_{1,2}x_2(t) + \cdots + a_{1,m}x_m(t)$$
$$\& \; x_2(t+1) = a_{2,1}x_1(t) + a_{2,2}x_2(t) + \cdots + a_{2,m}x_m(t)$$
$$\& \ldots \Rightarrow Ax(t) = x(t+1) \Rightarrow x(t) = A^t x(0)$$
- Finding the state of x at an arbitrary time t would be tedious since it would require t multiplications of A. However there is a simpler way…
- **Solving Dynamical Systems**
  - Step 1: Diagonalize $A \to A = SDS^{-1}$
    - Find an eigenbasis $\mathfrak{B}$ for A such that $\mathfrak{B} = \{v_1, \ldots, v_n\}$. Then,
    $$S = \begin{bmatrix} | & | & | & | \\ v_1 & v_2 & \ldots & v_n \\ | & | & | & | \end{bmatrix}, \quad D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$
    Then $\boldsymbol{A^t = SD^tS^{-1}}$
  - Step 2: Write *x(0)* as a linear combination of $v_1, v_2, \ldots, v_n$ (eigenvectors)
    - $S[\vec{x}]_{\mathfrak{B}} = x(0) \to [\vec{x}]_{\mathfrak{B}} = S^{-1}\vec{x} = [c_1, \; c_2, \ldots, c_n]^T$
    - $\vec{x}(0) = c_1v_1 + c_2v_2 + \cdots + c_nv_n$
  - Step 3: Rewrite $A^tx(0)$
    - $A^tx(0) = A^tc_1v_1 + \cdots + A^tc_nv_n = \lambda_1c_1v_1 + \cdots + \lambda_nc_nv_n$

Calculating the Determinant

- (*see Properties of Determinant/ Row Operations*)
- Cofactor Expansion (Laplace Expansion):
  - $$|B| = b_{i1}C_{i1} + b_{i2}C_{i2} + \cdots + b_{in}C_{in} = b_{1j}C_{1j} + b_{2j}C_{2j} + \cdots + b_{nj}C_{nj} = \sum_{j'=1}^{n} b_{ij'}C_{ij'} = \sum_{i'=1}^{n} b_{i'j}C_{i'j}$$
  where $C_{ij} = (-1)^{i+j}M_{ij}$, and b_{ij} is the value excluded by finding the minor M_{ij} for the cofactor C_{ij}

Subspace Given by Equation

- V is the subspace given by the equation:
$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = 0$$
  - Finding the kernel of V:
    - $$c_1x_1 + c_2x_2 + \cdots + c_nx_n = \begin{bmatrix} c_1 & c_2 & \ldots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0$$
    $$\Rightarrow M = \begin{bmatrix} c_1 & c_2 & \ldots & c_n & \vdots & 0 \end{bmatrix}$$
    $$c_1x_1 = -c_2x_2 - \cdots - c_nx_n$$
    $$x_2 = r$$
    $$x_3 = s$$
    $$\ldots$$
    $$x_n = z$$
    Express $\vec{x}$ as a linear combination.
  - Finding the matrix N such that V = Im(N):
    - Put the vectors of the linear combination from the result above into a matrix.

Rank

- rank(A) = rk(A) = # of leading 1s in rref(A)
- Also represents the number of linearly independent columns of A, a.k.a. the *dimension* of the vector space generated by A's columns
- **Properties**
  - rank(A) ≤ n and rank(A) ≤ m => rank(A) ≤ min(n, m)
  - If A is inconsistent, then rk(A) < n since at least one row will consist of all zeros
  - **If A has a unique solution, then rank(A) = m, since every variable must be bounded to a value (non-free)**

Properties of Transpose

- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^TA^T$
- $rank(A^T) = rank(A)$
- $(A^{-1})^T = (A^T)^{-1}$

Quadratic Forms

- A function $q(x_1, \; x_2, \ldots, x_n)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is called a *quadratic form* if it is a linear combination of the form x_ix_j (where i and j may be equal).
- A quadratic form can be written as $q(\vec{x}) = \vec{x} \cdot A\vec{x} = \vec{x}^TA\vec{x}$, where A is symmetric.
- The matrix A of q is given by the following rules:
  - $a_{ii} =$ the coefficient of x_i^2
  - $a_{ij} = a_{ji} = \frac{1}{2}$ (the coefficient of x_ix_j)
- **Diagonalizing a Quadratic Form**
  - Given $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, let $\mathfrak{B}$ be an orthonormal eigenbasis for A (*see symmetric matrices),* with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Then, $q(\vec{x}) = \lambda_1{c_1}^2 + \cdots + \lambda_n{c_n}^2$, where c_i are the coordinates of x with respect to $\mathfrak{B}$.
- **Definiteness of a Quadratic Form**
  - Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, we say that:
    - A is *positive definite* if q(x) > 0 for all x in $\mathbb{R}^n$
    - A is *positive semi-definite* if q(x) ≥ 0 for all x in $\mathbb{R}^n$
    - *Negative definite* and *negative semi-definite* are defined analogously.
    - If q(x) takes on both positive and negative values, then it is said to be indefinite
  - A symmetric matrix A is positive definite iff all of its eigenvalues are positive. The matrix A is positive semi-definite iff all of its eigenvalues are positive or zero. It works analogously for negative definite and negative semi-definite. Finally, the matrix A is indefinite if it has both positive and negative eigenvalues.

Singular Value Decomposition

- The *singular values* of a matrix A are the square roots of the eigenvalues of the matrix A^TA. It is customary to denote them in decreasing order:
$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$
  - If A is a $n \times m$ matrix of rank r, then the singular values $\sigma_1, \sigma_2, \ldots, \sigma_r$ are non-zero, while $\sigma_{r+1}, \ldots, \sigma_m$ are zero.
- A can be decomposed using *singular value decomposition*, such that $AV = U\Sigma \Rightarrow A = U\Sigma V^T$, where V is an orthogonal matrix formed by the orthonormal eigenbasis of A^TA. Then
$$U = \begin{bmatrix} | & | & | & | & | & | & | \\ u_1 & u_2 & \ldots & u_r & 0 & \ldots & 0 \\ | & | & | & | & | & | & | \end{bmatrix}$$
where $u_1 = \frac{Av_1}{\sigma_1}, u_2 = \frac{Av_2}{\sigma_2}, \ldots, u_r = \frac{Av_r}{\sigma_r}$, and
$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \vdots \\ & 0 & \cdots & 0 \end{bmatrix}$$
  - Note: If A has a singular value of zero, it's corresponding vector should be placed last in *V*, it should not have a corresponding vector in *U*, and Σ should include an additional column but not row for the singular value.