Orthogonal Matrices

$Q_{n\times n}$ is an orthogonal matrix iff has orthonormal cols:
Properties:

1. $(Q\vec{x})^T(Q\vec{z}) = \vec{x}^T z$ (angle preserving)
2. $||Q\vec{x}|| = ||\vec{x}||$ (length preserving)
3. $Q^T = Q^{-1}$ (true iff Q is orthogonal)
4. $\det(Q) = \pm 1$
5. Product of orthogonal matrices is orthogonal

$\det(Q) = 1$ for rotation $\det(Q) = -1$ for reflection

Diagonalization: $A = PDP^{-1}$

$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$, $P = \begin{bmatrix} \vec{\text{eigen}}_{\lambda_1} & \dots & \vec{\text{eigen}}_{\lambda_n} \end{bmatrix}$

A and D are similar $\Leftrightarrow A = PDP^{-1}$.
•A is similar to itself.
•A diagonalizable $\Leftrightarrow A^n = PD^nP^{-1}$
A **is diagonalizable iff** (equivalent statements):
• A has n linearly independent eigenvectors
• $\sum_i \text{GM}(\lambda_i) = n$. $\text{GM}(\lambda) :=$ number of eigenvectors for eigenvalue λ

Dot Product

• $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = (u_1)(v_1) + (u_2)(v_2)$
•$\vec{u}\cdot\vec{v} = \vec{u}^T\vec{v}$
•$\vec{u}\cdot\vec{u} = (||\vec{u}||)^2$
•$\vec{u}\cdot\vec{v} = 0 \Leftrightarrow \vec{u}, \vec{v}$ orthogonal $\Leftrightarrow ||\vec{u}|| + ||\vec{v}|| = ||\vec{u}+\vec{v}||$
•$(\vec{a}+\vec{b})\cdot(\vec{c}+\vec{d}) = \vec{a}\cdot\vec{c} + \vec{a}\cdot\vec{d} + \vec{b}\cdot\vec{c} + \vec{b}\cdot\vec{d}$
•$|| \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T || = \sqrt{v_1^2 + v_2^2}$

Eigenvalues and Eigenvectors

Eigenvalues of triangular matrix along the diagonal.
• $\text{AM}(\lambda = 0) = \text{nullity}(A)$

To **Find Eigenvalues** of $A = \begin{bmatrix} -5 & 2 \\ -9 & 6 \end{bmatrix}$

Solve $|A - \lambda I_n| = 0 \Leftrightarrow \begin{vmatrix} -5-\lambda & 2 \\ -9 & 6-\lambda \end{vmatrix} = 0$

To **Find Eigenvectors for** λ_i, solve:

$\left[\begin{array}{cc:c} -5-\lambda_i & 2 & 0 \\ -9 & 6-\lambda_i & 0 \end{array}\right]$

Linear Programming

Canonical formulation: Maximize $f(\vec{x}) = \vec{c}^T\vec{x}$
Subject to $A\vec{x} \le \vec{b}; \vec{x}_i \ge 0 \forall i$
Solution: Feasible intersection point which maximizes objective function

Orthogonal Compliments

Definition: if W is a subspace of $\mathbb{R}^n$, then $W^\perp$ contains all vectors $\in \mathbb{R}^n$ perpendicular to W

Properties:

1. $W^\perp = \text{Null}(A^T)$
2. $\dim(W) + \dim(W^\perp) = n$
3. $(W^\perp)^\perp = W$
4. $W \cap W^\perp = \{\vec{0}\}$
5. $(\mathbb{R}^n)^\perp = \{\vec{0}\}$; $\{\vec{0}\}^\perp = \mathbb{R}^n$

Subspaces of $\mathbb{R}^n$:

• $\text{Row}(A) = (\text{Nul}(A))^\perp$
• $\text{Nul}(A) = (\text{Row}(A))^\perp$
• $\dim(\text{Row}(A)) = \text{Rank}(A)$
•$\dim(\text{Nul}(A)) = n$ - $\text{Rank}(A)$

Subspaces of $\mathbb{R}^m$:

• $\text{Col}(A) = (\text{Nul}(A^T))^\perp$
•$\text{Nul}(A^T) = (\text{Col}(A))^\perp$
•$\dim(\text{Col}(A)) = \text{Rank}(A)$
•$\dim(\text{Nul}(A^T)) = m$ - $\text{Rank}(A)$

Example: find a basis for Row(A)

$\vec{v} = \begin{bmatrix} 1 & 2 & 1 & 2 \end{bmatrix}^T$, and $A\vec{v} = \vec{0}$
$A\vec{v} = \vec{0}$, so $\vec{v} \in \text{Nul}(A)$.
Let W =Nul(A), so $W^\perp = \text{Row}(A)$
$\Rightarrow \text{Nul}(\vec{v}^T)$ is a basis for Row(A).

$$\text{Nul}(\begin{bmatrix} 1 & 2 & 1 & 2 \end{bmatrix}) = \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

= basis for Row(A)

Span

Def The span of a set of vectors is the set of linear combinations of the vectors.
Check if $\vec{w} \in \text{span}\{\vec{u}, \vec{v}\}$:

$\left[\begin{array}{cc:c} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \end{array}\right]$ $\vec{w} \in \text{span}\{\vec{u}, \vec{v}\}$ iff that system has a solution.

Check if $\text{span}\{\vec{u}, \vec{v}, \vec{w}\} = \mathbb{R}^3$:
(equivalent to $\text{col}[\vec{u}, \vec{v}, \vec{w}] = \mathbb{R}^3$)

$\begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$ True iff $\exists$ a pivot in each row/column

Orthogonal Decomposition

W a subspace of $\mathbb{R}^n \Leftrightarrow y = \hat{y} + z$, s.t. $\hat{y} \in W, z \in W^\perp$
$\hat{y} = proj_{u_1} y + \ldots + proj_{u_p} y = proj_W y = \sum_{i=1}^p \frac{y\cdot u_i}{||u_i||^2}$

Matrix Properties

$A_{m\times n} \Rightarrow$ transformation $\vec{x} \to A\vec{x}$ is from $\mathbb{R}^n \to \mathbb{R}^m$
$\text{Rank}(A) + \text{nullity}(A) = n$
A^TA & AA^T:

1. have same non 0 eigenvalues
2. Symmetric, positive semidefinite

•A^TA invertible iff A has independent columns

Determinants	$\det(AB) = \det(A)\det(B)$ $\det(A^{-1}) = \frac{1}{\det(A)}$ $\det(\text{adj}(A)) = \det(A)^{n-1}$ $\det(A^n) = (\det(A))^n$ $\det(cA) = c^n \det(A)$ $\det(A^T) = \det(A) = \lvert A\rvert$ $\det(A_{\text{tringl}}) = \prod_{i=1}^n A_{(i,i)}$ $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$
Inverse	$(AB)^{-1} = B^{-1}A^{-1}$ $(A^{-1})^{-1} = A$ $(kA^{-1}) = \frac{1}{k}A$ $(A^{-1})^{-1} = I_n$ $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
Transpose	$(A^T)^T = A$ $(AB)^T = B^TA^T$ $(kA^T) = kA^T$ $(A+B)^T = A^T + B^T$

Injectivity and Surjectivity

Injective: (one-to-one)

1. $f(x) = f(y) \leftarrow x = y$ (map to the same point)
2. $\text{rank}(A) = \text{n}$
3. pivot in every column

Surjective: (onto)

1. $\forall \vec{b} \in \mathbb{R}^m, \exists \vec{x} \in \mathbb{R}^n$ such that $f(\vec{x}) = \vec{b}$
2. $\text{rank}(A) = \text{m}$
3. pivot in every row

Gram-Schmidt Process

Gram-Schmidt Process: (orthogonalize basis)
$\vec{v_1} = \vec{x_1}$ (remember to normalize at end)
$\vec{v_2} = \vec{x_2} - \frac{\vec{x_2}\cdot\vec{v_1}}{\vec{v_1}\cdot\vec{v_1}}\vec{v_1}$
$\vec{v_p} = \vec{x_p} - \frac{\vec{x_p}\cdot\vec{v_1}}{\vec{v_1}\cdot\vec{v_1}}\vec{v_1} - \frac{\vec{x_p}\cdot\vec{v_2}}{\vec{v_2}\cdot\vec{v_2}}\vec{v_2} - \ldots \frac{\vec{x_p}\cdot\vec{v_{p-1}}}{\vec{v_{p-1}}\cdot\vec{v_{p-1}}}\vec{v_{p-1}}$
QR Factorization ($A = QR$):
$A_{n\times q}$ has independent columns. Apply Gram-Schmidt on columns to get $\{\vec{v_1}, .., \vec{v_q}\}$, then normalize to $\{\vec{u_1}, .., \vec{u_q}\}$
•$Q = [\vec{u_1}, .., \vec{u_q}]$
•$R = Q^TA$

Invertible Matrix Theorem

1. A is row-equivalent to the $n \times n$ I_n.
2. A has n pivot positions.
3. The equation $Ax = 0$ has only the trivial solution.
4. The columns of A form a linearly independent set.
5. The linear transformation $x \mapsto Ax$ is one-to-one.
6. $\forall \vec{b} \in \mathbb{R}^n$, $A\vec{x} = \vec{b}$ has a unique solution.
7. The columns of A span $\mathbb{R}^n$.
8. The linear transformation $x \mapsto Ax$ is a surjection.
9. There is an $n \times n$ matrix C such that $CA = I_n$.
10. There is an $n \times n$ matrix D such that $AD = I_n$.
11. The transpose matrix A^T is invertible.
12. The columns of A form a basis for $\mathbb{R}^n$.
13. The column space of A is equal to $\mathbb{R}^n$.
14. The dimension of the column space of A is n.
15. The rank of A is n.
16. The null space of A is $\{0\}$.
17. The dimension of the null space of A is 0.
18. 0 fails to be an eigenvalue of A.
19. $\det(A) \neq 0$
20. The orthogonal complement of $\text{Col}(A)$ is $\{0\}$.
21. The orthogonal complement of $\text{Nul}(A)$ is $\mathbb{R}^n$.
22. The row space of A is $\mathbb{R}^n$.

Projections

1. Projection matrices are symmetric.
2. $P^2 = P$ (1&2 $\Leftrightarrow$ P is projection matrix)

•P projects onto $\text{Col}(P)$
•$proj_{\vec{u}}\vec{v} = \frac{\vec{v} \cdot \vec{u}}{||\vec{u}||^2}\vec{u}$
•$proj_w\vec{b} = P\vec{b}$, $w = \text{col}(A)$
•$P_{\text{Row}(A)} = I - P_{\text{Col}(A)}$
•$\vec{v} \in \text{Nul}(A) \Rightarrow \forall \vec{u}$ s.t. $\vec{u} \cdot \vec{v} = 0, \vec{u} \in \text{Row}(A)$
•$\vec{v} \in w \Leftrightarrow proj_w\vec{v} = \vec{v}$, so $P\vec{v} = \vec{v} \Rightarrow v \in w$
•$\vec{v} \in w^\perp \Leftrightarrow proj_w\vec{v} = 0$, so $P\vec{v} = 0 \Rightarrow v \in w^\perp$
• $\text{Rank}(P_{\text{subspace}}) = \dim(\text{subspace})$
•$||P_{\text{Row}(A)}\vec{x}||$ is shortest distance from $\vec{x}$ to $\text{Nul}(A)$
•$P = A(A^TA)^{-1}A^T$ In general (subspace)
•$P = \frac{(\vec{a})(\vec{a}^T)}{(\vec{a}^T)(\vec{a})}$ (for line only)

Shortcut for Orthonormal vectors:
$P\vec{b} = (\vec{a_1} \cdot \vec{b})\vec{a_1} + (\vec{a_2} \cdot \vec{b})\vec{a_2} + ... + (\vec{a_n} \cdot \vec{b})\vec{a_n}$

Shortcut for Orthogonal vectors:
$P\vec{b} = proj_{\vec{a_1}}b + proj_{\vec{a_2}}b + ... + proj_{\vec{a_n}}b$

Example: $P_W \begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\-2\\-2\end{bmatrix}$, find W

$\vec{b} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}^T \notin W, \vec{b} \notin W^\perp$

so $\vec{z} = \vec{b} - P\vec{b}$, and $\vec{z} \in W^\perp$

$\vec{z} = \begin{bmatrix}1\\1\\1\end{bmatrix} - \begin{bmatrix}0\\-2\\-2\end{bmatrix} = \begin{bmatrix}1\\3\\3\end{bmatrix} \Rightarrow W$ given by $x+3y+3z = 0$

Singular Value Decomposition

$A = U\Sigma V^T$; orthogonal, diagonal, orthogonal

Finding an SVD: $A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n}$

1. orthogonally diagonalize A^TA
2. Σ: diagonal decreasing $\sqrt{\text{eigenvalues}}$, 0's elsewhere
3. V: eigen's corresponding to Σ's $\sqrt{\text{eigenvalues}}$

4a. $\vec{u_i} = \frac{1}{\sigma_i}A\vec{v_i}$, while possible

4b. $\vec{u}_{(n+1...m)} = \text{Nul}\left(\begin{bmatrix}\vec{u_1}^T\\ \vdots \\ \vec{u_n}^T\end{bmatrix}\right)$ ($\because U$ is orthogonal)

Pieces of SVD: $A = \sum_{i=1}^{r} = \sigma_i\vec{u}_i\vec{v}_i^T$
• $\sum_{i=1}^{\phi} = \sigma_i\vec{u}_i\vec{v}_i^T$ is best rank ϕ approximation of A

Reduced SVD: $A = U_r\Sigma_rV_r^T$
•$A_{m\times n} = U_{m\times r}\Sigma_{r\times r}V^T_{r\times n}$

Psuedo-Inverse $A^+ = V_r\Sigma_r^{-1}U_r^T$**:**
•$\hat{x} = A^+\vec{b}$
•$AA^+ = P_{\text{Col}(A)} = U_rU_r^T$
•$A^+A = P_{\text{Row}(A)} = V_rV_r^T$

Geometry of SVD:

1. σ_1: max length of ellipsoid
2. $\vec{v_1}$: $A\vec{v}$ maximizes len(ellipsoid)
3. $\vec{u_1}$: direction of largest axis

• rotate, stretch, rotate (like orthog. diagonalization)
•$A\text{Row}(A) = \text{Col}(A) = \text{Row}(A^T)$
• $||\vec{x}|| = 1 \Rightarrow \max\ (||A\vec{x}||) = \sigma_1$. Occurs at $\vec{x}$ = corresponding eigen.

Polar Decomposition: $A = QS$: (rotate, stretch)

1. $S = V\Sigma V^T$ is symmetric positive semidefinite
2. $Q = UV^T$ is orthogonal

Bases of 4 fundamental subspaces from SVD:

1. $\text{Col}(A)$: first r cols of U
2. $\text{Nul}(A^T)$: cols beyond rth of $U; (r+1...m)$
3. $\text{Row}(A)$: first r rows of V^T
4. $\text{Nul}(A)$: rows beyond rth of $V^T; (r+1...n)$

Applications of SVD to Stats

$B_{m\times n}$: A w/ elements as value - row's mean

Covariance Matrix: $S_{m\times m} = \frac{1}{n-1}BB^T$

PCA: Principal Components are unit eigen of S
• Total Variance $= \sum_{i=1}^{m} S_{(i,i)} = \sum_{i=1}^{m}\lambda_i$
• var(principle component i) $= \lambda_i$

Least Squares

$\hat{x} = (A^TA)^{-1}A^T\vec{b}$

- **Normal equations** for $A\vec{x} = \vec{b}$: $A^TA\vec{x} = A^T\vec{b}$

•$A^TA\vec{x} = A^T\vec{b}$ has same solutions as $\hat{x}$.
•$A\hat{x} = proj_w\vec{b}$
•$\hat{x} = R^{-1}Q^T\vec{b}$, where $A = QR$ factorization
•$\hat{x} = \vec{x}$ s.t. $A\hat{x}$ is as close as possible to $\vec{b}$
•$\hat{x} = A^+\vec{b}$

Orthogonal Diagonalization & Spec. Decomp.

- $A = QDQ^T$ for orthogonal Q
- $A^T = A \Rightarrow A\vec{x} = (A\vec{x})^T$

Spectal Theorem for Orthognal Matrices:

1. A is orthogonally diagonalizable iff $A = A^T$
2. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation
3. $\vec{Eigen}$ for different λ's orthogonal
4. Symmetric A has real eigenvalues

Spectral Decomposition:
$A = \sum_{i=1}^{n}\lambda_i\vec{u_i}\vec{u_i}^T \Leftrightarrow (Q = [\vec{u_1}, \vec{u_2}, ..., \vec{u_n}])$

Quadratic Forms of Symmetric Matrices

$A_{n\times n}$ has quadratic form $Q(\vec{x}) = \vec{x}^TA\vec{x}$ for $\vec{x} \in \mathbb{R}^n$

$Q(x) = a_1x_1^2 + a_2x_1x_2 + a_3x_2^2 \Leftrightarrow A = \begin{bmatrix}a_1 & \frac{a_2}{2}\\ \frac{a_2}{2} & a_2\end{bmatrix}$

- Definite: $Q(\vec{x}) < 0\forall x \neq 0 \Leftrightarrow \lambda < 0\forall\lambda$
- Semidefinite: $Q(\vec{x}) \leq 0\forall x \neq 0 \Leftrightarrow \lambda \leq 0\forall\lambda$
+ Definite: $Q(\vec{x}) > 0\forall x \neq 0 \Leftrightarrow \lambda > 0\forall\lambda$
+ Semidefinite: $Q(\vec{x}) \geq 0\forall x \neq 0 \Leftrightarrow \lambda \geq 0\forall\lambda$
• Semidefinite has ∞ mins/maxes
• Definite matrices are symmetric
• $Q(x)$ of $A_{n\times n}$ has n x_i^2 terms
• $||\vec{x}|| = 1 \Rightarrow \min Q(\vec{x}) = \min(\lambda)$, $\max Q(\vec{x}) = \max(\lambda)$
• Max occurs at $\vec{x} = \frac{1}{||\text{eigen}||}$eigen for $\max(\lambda)$

Principal Axes Theorem:
Symmetric A has orthogonal change of vars $\vec{x} = P\vec{y}$, transforms $\vec{x}^TA\vec{x}$ into $\vec{y}^TQ\vec{y}$ w/ no cross-prod term.
- After change of vars, coeff's on y_i^2 are eigenvalues