

Discussion 1: Invertibility of $A^\top A$ and Eigenvalues

1. Review: Key Concepts

Before diving into the problems, we review the essential definitions and properties.

1.1 Null Space

Definition 1 (Null Space). The **null space** (or kernel) of a matrix $A \in \mathbb{R}^{m \times n}$ is the set of all vectors that A maps to zero:

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$$

The null space is always a subspace of \mathbb{R}^n . It captures information about what A “loses”: if $x \in \mathcal{N}(A)$, then A cannot distinguish x from the zero vector.

1.2 Full Column Rank

Definition 2 (Full Column Rank). A matrix $A \in \mathbb{R}^{m \times n}$ has **full column rank** if $\text{rank}(A) = n$, meaning all n columns are linearly independent.

Equivalent Conditions for Full Column Rank:

1. $\text{rank}(A) = n$
2. $\mathcal{N}(A) = \{0\}$ (only the zero vector maps to zero)
3. The columns of A are linearly independent
4. $A^\top A \in \mathbb{R}^{n \times n}$ is invertible

Note: Full column rank requires $m \geq n$ (at least as many rows as columns).

1.3 Invertibility of Square Matrices

A square matrix $M \in \mathbb{R}^{n \times n}$ is invertible if and only if $\mathcal{N}(M) = \{0\}$.

Why? If $Mx = 0$ has only the trivial solution $x = 0$, then M has full rank, so it is invertible. Conversely, if M is invertible and $Mx = 0$, multiplying both sides by M^{-1} gives $x = 0$.

1.4 Eigendecomposition

Definition 3 (Eigenvalue and Eigenvector). For a square matrix $A \in \mathbb{R}^{n \times n}$, a scalar λ is an **eigenvalue** and a nonzero vector v is a corresponding **eigenvector** if:

$$Av = \lambda v$$

Definition 4 (Diagonalizable Matrix). A matrix $A \in \mathbb{R}^{n \times n}$ is **diagonalizable** if there exists an invertible matrix P and a diagonal matrix Λ such that:

$$A = P\Lambda P^{-1}$$

Here, the columns of P are eigenvectors of A , and the diagonal entries of Λ are the corresponding eigenvalues.

Eigendecomposition: If $A = P\Lambda P^{-1}$ where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, then:

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

1.5 Determinant Properties

Key Determinant Facts:

1. $\det(XY) = \det(X)\det(Y)$ for square matrices X, Y
2. $\det(X^{-1}) = \frac{1}{\det(X)}$ for invertible X
3. For a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$:

$$\det(\Lambda) = \prod_{i=1}^n \lambda_i$$

2. Problem 1: Invertibility of $A^\top A$

Let $A \in \mathbb{R}^{m \times n}$. We will show that A has full column rank if and only if $A^\top A$ is invertible.

2.1 Part (a): $\mathcal{N}(A) \subseteq \mathcal{N}(A^\top A)$

Goal: Show that if $x \in \mathcal{N}(A)$, then $x \in \mathcal{N}(A^\top A)$.

Step 1: Start with the hypothesis.

Suppose $x \in \mathcal{N}(A)$. By definition of the null space, this means:

$$Ax = 0$$

Step 2: Apply A^\top to both sides.

Multiplying both sides of $Ax = 0$ by A^\top :

$$A^\top(Ax) = A^\top \cdot 0 = 0$$

Step 3: Recognize the left-hand side.

By associativity of matrix multiplication:

$$(A^\top A)x = 0$$

Step 4: Conclude.

This shows $x \in \mathcal{N}(A^\top A)$. Therefore, $\mathcal{N}(A) \subseteq \mathcal{N}(A^\top A)$. \square

Intuition

If A already kills x (sends it to zero), then applying A^\top afterward still gives zero. Anything in the null space of A is automatically in the null space of $A^\top A$.

2.2 Part (b): $\mathcal{N}(A^\top A) \subseteq \mathcal{N}(A)$

Goal: Show that if $x \in \mathcal{N}(A^\top A)$, then $x \in \mathcal{N}(A)$.

Step 1: Start with the hypothesis.

Suppose $x \in \mathcal{N}(A^\top A)$. By definition:

$$A^\top Ax = 0$$

Step 2: Take the inner product with x .

Multiply both sides on the left by x^\top :

$$x^\top A^\top Ax = x^\top \cdot 0 = 0$$

Step 3: Recognize the quadratic form.

The left-hand side can be rewritten using the property $(Ax)^\top = x^\top A^\top$:

$$x^\top A^\top Ax = (Ax)^\top (Ax) = \|Ax\|^2$$

Step 4: Apply positive definiteness of norms.

We now have:

$$\|Ax\|^2 = 0$$

Since the squared norm of a vector equals zero if and only if the vector itself is zero:

$$Ax = 0$$

Step 5: Conclude.

This shows $x \in \mathcal{N}(A)$. Therefore, $\mathcal{N}(A^\top A) \subseteq \mathcal{N}(A)$. \square

The Norm Trick

This is a fundamental technique: to show $Ax = 0$, show that $\|Ax\|^2 = 0$. The “ $x^\top(\cdot)x$ ” sandwich turns the equation into a norm, and norms are zero only for the zero vector.

2.3 Combining Parts (a) and (b)

From parts (a) and (b), we have:

$$\mathcal{N}(A) \subseteq \mathcal{N}(A^\top A) \quad \text{and} \quad \mathcal{N}(A^\top A) \subseteq \mathcal{N}(A)$$

By double inclusion, we conclude:

$$\mathcal{N}(A) = \mathcal{N}(A^\top A)$$

The null spaces of A and $A^\top A$ are identical.

2.4 Part (c): Full Column Rank Implies $A^\top A$ Invertible

Goal: Show that if A has full column rank, then $A^\top A$ is invertible.

Step 1: Translate full column rank.

A has full column rank means $\text{rank}(A) = n$. By the rank-nullity theorem:

$$\dim(\mathcal{N}(A)) = n - \text{rank}(A) = n - n = 0$$

Therefore, $\mathcal{N}(A) = \{0\}$.

Step 2: Apply the null space equality.

From the result above, $\mathcal{N}(A) = \mathcal{N}(A^\top A)$. Since $\mathcal{N}(A) = \{0\}$:

$$\mathcal{N}(A^\top A) = \{0\}$$

Step 3: Conclude invertibility.

The matrix $A^\top A \in \mathbb{R}^{n \times n}$ is a square matrix with trivial null space. A square matrix is invertible if and only if its null space is $\{0\}$. Therefore, $A^\top A$ is invertible. \square

Main Result: For $A \in \mathbb{R}^{m \times n}$:

$$A \text{ has full column rank} \iff A^\top A \text{ is invertible}$$

Why This Matters

This result is fundamental to least squares. When solving the normal equations $A^\top A x = A^\top b$, we need $A^\top A$ to be invertible to get a unique solution $x = (A^\top A)^{-1} A^\top b$. Full column rank of A guarantees this.

3. Problem 2: Eigenvalues and Determinants

Let $A \in \mathbb{R}^{n \times n}$ be a diagonalizable matrix with eigendecomposition $A = P\Lambda P^{-1}$, where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$.

3.1 Part (a): Powers of Diagonalizable Matrices

Goal: Show that $A^m = P\Lambda^m P^{-1}$ for any positive integer m .

Step 1: Compute A^2 .

Starting from $A = P\Lambda P^{-1}$:

$$\begin{aligned} A^2 &= A \cdot A \\ &= (P\Lambda P^{-1})(P\Lambda P^{-1}) \\ &= P\Lambda(P^{-1}P)\Lambda P^{-1} \\ &= P\Lambda I\Lambda P^{-1} \\ &= P\Lambda^2 P^{-1} \end{aligned}$$

The key step is that $P^{-1}P = I$ cancels in the middle.

Step 2: Compute A^3 .

$$\begin{aligned} A^3 &= A^2 \cdot A \\ &= (P\Lambda^2 P^{-1})(P\Lambda P^{-1}) \\ &= P\Lambda^2(P^{-1}P)\Lambda P^{-1} \\ &= P\Lambda^3 P^{-1} \end{aligned}$$

Step 3: General pattern by induction.

Base case: For $m = 1$, $A^1 = P\Lambda^1 P^{-1}$ holds by definition.

Inductive step: Assume $A^k = P\Lambda^k P^{-1}$ for some $k \geq 1$. Then:

$$\begin{aligned} A^{k+1} &= A^k \cdot A \\ &= (P\Lambda^k P^{-1})(P\Lambda P^{-1}) \\ &= P\Lambda^k(P^{-1}P)\Lambda P^{-1} \\ &= P\Lambda^{k+1} P^{-1} \end{aligned}$$

By induction, the formula holds for all positive integers m .

Powers of Diagonalizable Matrices:

$$A^m = P\Lambda^m P^{-1}$$

where $\Lambda^m = \text{diag}(\lambda_1^m, \dots, \lambda_n^m)$.

Why Diagonalization is Powerful

Computing A^{100} directly requires 99 matrix multiplications. With diagonalization, we compute Λ^{100} (just raise each diagonal entry to the 100th power) and do two matrix multiplications: $P\Lambda^{100}P^{-1}$. This is vastly more efficient.

3.2 Part (b): Determinant Equals Product of Eigenvalues

Goal: Show that $\det(A) = \prod_{i=1}^n \lambda_i$.

Step 1: Apply determinant properties to the eigendecomposition.

Starting from $A = P\Lambda P^{-1}$, take determinants of both sides:

$$\det(A) = \det(P\Lambda P^{-1})$$

Step 2: Use the multiplicative property.

The determinant of a product equals the product of determinants:

$$\det(P\Lambda P^{-1}) = \det(P) \cdot \det(\Lambda) \cdot \det(P^{-1})$$

Step 3: Simplify using $\det(P^{-1}) = \frac{1}{\det(P)}$.

$$\det(P) \cdot \det(\Lambda) \cdot \det(P^{-1}) = \det(P) \cdot \det(\Lambda) \cdot \frac{1}{\det(P)} = \det(\Lambda)$$

The $\det(P)$ terms cancel.

Step 4: Compute the determinant of a diagonal matrix.

For a diagonal matrix, the determinant is the product of diagonal entries:

$$\det(\Lambda) = \det \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \lambda_1 \cdot \lambda_2 \cdots \lambda_n = \prod_{i=1}^n \lambda_i$$

Step 5: Conclude.

$$\det(A) = \prod_{i=1}^n \lambda_i \quad \square$$

Determinant and Eigenvalues:

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$$

The determinant of a matrix equals the product of its eigenvalues (counted with multiplicity).

Corollary: Invertibility Test

Since $\det(A) = \prod \lambda_i$, a matrix is invertible (i.e., $\det(A) \neq 0$) if and only if none of its eigenvalues are zero. Equivalently: A is singular $\iff 0$ is an eigenvalue of A .

Note on Generality

The result $\det(A) = \prod \lambda_i$ holds for *all* square matrices, not just diagonalizable ones. The proof for non-diagonalizable matrices uses the characteristic polynomial: $\det(A - \lambda I) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$. Setting $\lambda = 0$ gives $\det(A) = \prod \lambda_i$.

Key Takeaways

1. **Null space equality:** $\mathcal{N}(A) = \mathcal{N}(A^\top A)$. The “norm trick” ($x^\top A^\top A x = \|Ax\|^2$) is essential for proving one direction.
2. **Full column rank criterion:** A has full column rank $\iff A^\top A$ is invertible. This underlies uniqueness in least squares.
3. **Diagonalization simplifies powers:** $A^m = P \Lambda^m P^{-1}$. The $P^{-1}P = I$ cancellation is the key mechanism.
4. **Determinant via eigenvalues:** $\det(A) = \prod \lambda_i$. Zero eigenvalue \iff singular matrix.