MA-GY 7043: Linear Algebra II

Moore-Penrose Quasi-Inverse

Deane Yang

Courant Institute of Mathematical Sciences
New York University

April 27 2026

Outline I

Moore-Penrose
Quasi-Inverse

Moore-Penrose Quasi-Inverse

System of Linear Equations

- Consider a system of n equations with m unknowns,

$$a_1^1 x^1 + \cdots + a_m^1 x^m = y^1$$
$$\vdots$$
$$a_1^n x^1 + \cdots + a_m^n x^m = y^n$$

- Usually, there is no solution
- And, even if there is a solution, it is usually not unique
- Basic examples
  - 1 equation in 1 unknown

$$3x = 1$$

  - 1 equation in 2 unknowns

$$x + y = 1$$

  - 2 equations in 2 unknowns

$$x + y = 1$$
$$x + y = 2$$

Matrix Equation

- Given $A \in \mathcal{M}_{n\times m}(\mathbb{F})$ and $y \in \mathbb{F}^n$, we want to solve for $x \in \mathbb{F}^m$ such that
$$Ax = y$$
- There is a solution if and only if $y \in \text{image}(A)$
- If a solution exists, then it is unique if and only if $\ker A = \{0\}$
- It is possible that $y \notin \text{image}(A)$, because A and y are from inexact measurements
- Instead, we look for best possible approximation

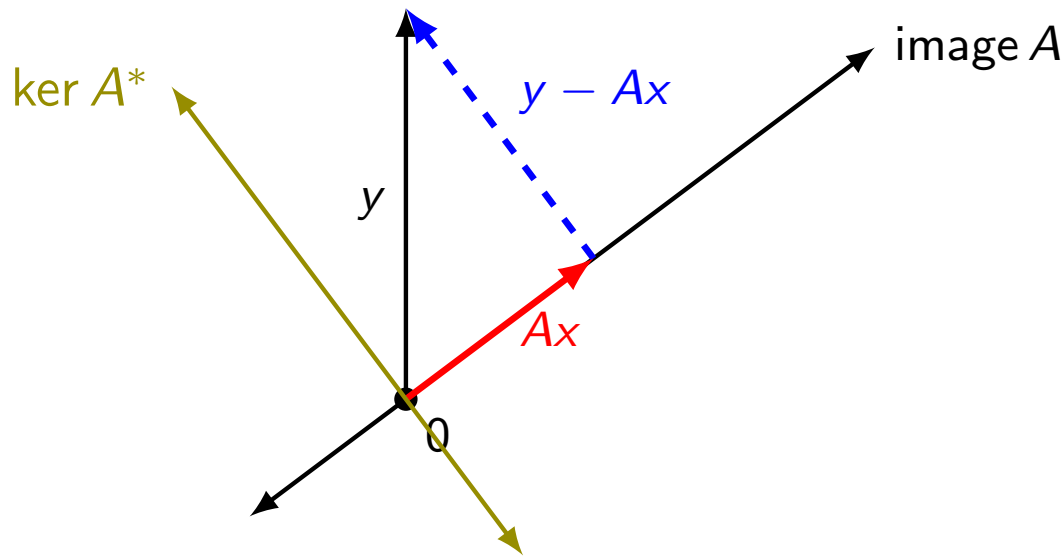
Quasi-Solution with Least Error

- Given $x \in \mathbb{F}^m$, define the error to be

$$\epsilon = L(x) - y \in \mathbb{F}^n$$

- Goal: Solve for x that minimizes the magnitude of the error, $\|\epsilon\|$
- An $x \in X$ that minimizes $\|\epsilon\|$ is called a **quasi-solution**
- If $y \in$ image L, then a quasi-solution is a solution
- A quasi-solution need not be unique

Geometric Perspective



- If Ax is closest to y, then
  - $y - Ax$ is orthogonal to image A
- Recall that $(\text{image } A)^\perp = \ker A^*$
  - $y \in (\text{image}(A))^\perp$ if and only if for every x,

$$0 = (A(x), y) = (x, A^*(y)),$$

which implies $A^*(y) = 0$, i.e., $y \ker(A^*)$

- Therefore, A is closest to y if

$$A^*(y - Ax)) = 0, \text{ i.e., } A^*Ax = A^*y$$

Example

- Consider the system of equations

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

- Equivalently,

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

- There is no solution

Quasi-Solution

- Let

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

- (x, y, z) is a quasi-solution if

$$A^* A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^* \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$

Quasi-Solution Via Row Reduction

- (x, y, z) is a quasi-solution if

Moore-Penrose
Quasi-Inverse

$$\begin{bmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$

$$\implies \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 0 \end{bmatrix}$$

$$\implies \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$$

$$\implies x + y = 2$$

$$z = 2$$

Quasi-Solution Error

- $\begin{bmatrix} x \\ 2-x \\ 2 \end{bmatrix}$ is a quasi-solution to $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$
- The error of the quasi-solution

$$\epsilon = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 2-x \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Error Comparison

- The error for any other (x, y, z) is

$$\begin{aligned}\epsilon &= \begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix} - \begin{bmatrix}3\\ 3\\ 3\end{bmatrix}\\ &= \begin{bmatrix}x+y+z-3\\ x+y-3\\ z-3\end{bmatrix}\\ &= \begin{bmatrix}1\\ -1\\ -1\end{bmatrix} + \begin{bmatrix}x+y+z-4\\ x+y-2\\ z-2\end{bmatrix}\end{aligned}$$

- The error magnitude squared is

$$\epsilon^2 = \left\|\begin{bmatrix}1\\ -1\\ -1\end{bmatrix}\right\|^2 + \left\|\begin{bmatrix}x+y+z-4\\ x+y-2\\ z-2\end{bmatrix}\right\|^2 \geq \left\|\begin{bmatrix}1\\ -1\\ -1\end{bmatrix}\right\|^2$$

Quasi-Solutions of $L(x) = y$

Moore-Penrose Quasi-Inverse

- Let X, Y be inner product spaces and $L : X \to Y$ be a linear map
- $L(x)$ is closest to y if

$$L^*L(x) = L^*(y),$$

- For any $y \in Y$, there is always a quasi-solution $x \in X$, i.e.,

$$\text{image}(L^*L) = \text{image}\, L^*,$$

- Geometrically, this is obvious
- Algebraically, use the singular value decomposition: If $L = V \circ \Sigma \circ W^*$, then

$$L^* \circ L = W \circ \Sigma^* \circ V^* \circ V \circ \Sigma \circ W^* = W \circ \Sigma^* \circ \Sigma \circ W^*,$$

  which is self-adjoint and $\Sigma^* \circ \Sigma$ is
- Alternatively
  - Recall that $\ker(L^*L) = \ker L$
  - Therefore, since L^*L is self-adjoint,

$$\text{image}(L^*L) = (\ker L^*L)^\perp = (\ker L)^\perp = \text{image}\, L^*$$

- If $v \in \ker L^*L = \ker L$, then $x + v$ is also a solution

Error Comparison

- A quasi-solution of the equation $L(x) = y$ satisfies

$$L^* L(x) = L^*(y)$$

  and therefore $L^*(L(x) - y) = 0$
- The error of the quasi-solution x is

$$\epsilon = L(x) - y$$

- The error of any $x' \in X$ is

$$\epsilon' = L(x') - y = L(x' - x) + L(x) - y = L(x' - x) + \epsilon$$

- On the other hand,

$$\begin{aligned}(L(x' - x), \epsilon) &= (x' - x, L^*(\epsilon)) \\ &= (x' - x, L^* L(x) - L^*(y)) \\ &= 0\end{aligned}$$

- Therefore, $\|\epsilon'\|^2 = \|\epsilon\|^2 + \|L(x' - x)\|^2$

Quasi-Solution when $L^*L : X \to X$ is Invertible

- If x is a quasi-solution, then

$$L^*L(x) = L^*(y)$$

- If the map $L^*L : X \to X$ is invertible, then the unique quasi-solution is

$$x = (L^*L)^{-1}L^*(y)$$

Solution with Minimal Magnitude

- Suppose $x \in X$ is a solution (not just a quasi-solution) of

$$Ax = y$$

- If $v \in \ker A$, then $x + v$ is also a solution,

$$A(x + v) = y$$

- There is a unique solution x with minimal magnitude

Minimal Magnitude Solution Via Orthogonal Projection

- For any $x' \in X$, there is a unique way to decompose x' into a sum
$$x' = x + (x' - x),$$
where $x \in (\ker A)^{\perp}$ and $x - x' \in \ker A$
- If x' is a solution to
$$Ax' = y,$$
then
$$Ax = A(x - x') + Ax' = y$$
- If $x_1, x_2 \in (\ker A)^{\perp}$ are both solutions, then
$$x_1 - x_2 \in (\ker A)^{\perp} \text{ and } x_1 - x_2 \in \ker A,$$
because
$$A(x_1 - x_2) = Ax_1 - Ax_2 = y - y = 0$$
Therefore, $x_1 - x_2 = 0$

Quasi-Solution with Minimal Magnitude

- A quasi-solution to
$$Ax = y$$
is a solution of
$$A^*Ax = A^*y$$
- There is a unique quasi-solution $x \in (\ker A^*A)^\perp = (\ker A)^\perp$

Example

- The quasi-solutions of the equation

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

  are

$$\begin{bmatrix} x \\ 2-x \\ 2 \end{bmatrix}, \ x \in \mathbb{F}$$

- The magnitude squared of each quasi-solution is

$$\left\| \begin{bmatrix} x \\ 2-x \\ 2 \end{bmatrix} \right\|^2 = x^2 + (2-x)^2 + 4 = 2((x-1)^2 + 3)$$

- The magnitude is minimized when $x = 1$ and therefore the Moore-Penrose quasi-solution is $(1, 1, 2)$

Moore-Penrose Quasi-Inverse Operator

- Let X and Y be inner product spaces and $L : X \to Y$ be a linear map
- There is a map $L^+ : Y \to X$ such that for any $y \in Y$, $x = L^+(y)$ is the unique quasi-solution with minimal magnitude of the equation
$$L(x) = y$$
- The map L^+ is called the **Moore-Penrose quasi-inverse** of L

Moore-Penrose Quasi-Inverse Operator

- The map
$$L|_{(\ker L)^\perp} : (\ker L)^\perp \to \text{image}\, L$$
is an isomorphism.
- Let
$$\pi : Y \to \text{image}\, L$$
be orthogonal projection
- The Moore-Penrose quasi-inverse operator is the map
$$L^+ : Y \to X,$$
given by
$$L^+(y) = \left(L|_{(\ker L)^\perp} \right)^{-1} (\pi(y)) \in (\ker L)^\perp \subset X$$

Quasi-Inverse of Diagonal Matrix

- Let $\Sigma : \mathbb{R}^m \to \mathbb{R}^m$ be the diagonal matrix such that for each $1 \leq k \leq m$,

$$\Sigma(\epsilon_k) = \begin{cases} s_k \epsilon_k & \text{if } 1 \leq k \leq r \\ 0 & \text{if } r+1 \leq k \leq m \end{cases}$$

- Therefore,

$$\Sigma(\epsilon_1 v^1 + \cdots + \epsilon_m v^m) = \epsilon_1 s_1 v^1 + \cdots + \epsilon_r s_r v^r$$

- The quasi-inverse of Σ satisfies the following:

$$\Sigma^+(\epsilon_1 v^1 + \cdots + \epsilon_m v^m) = \epsilon_1 s_1^{-1} v^1 + \cdots + \epsilon_r s_r v^r$$

- In particular,

$$\begin{aligned} \Sigma^+(\Sigma(\epsilon_1 v^1 + \cdots + \epsilon_m v^m)) &= \Sigma^+(\epsilon_1 s_1 v^1 + \cdots + \epsilon_r s_r v^r) \\ &= \epsilon_1 v^1 + \cdots + \epsilon_r v^r \\ &= \pi_r(v), \end{aligned}$$

where $\pi_r : \mathbb{R}^m \to \mathbb{R}^m$ is orthogonal projection onto the subspace spanned by $(\epsilon_1, \ldots, \epsilon_r)$

Quasi-Inverse Via Singular Value Decomposition

- Let the singular value decomposition of $L : X \to Y$ be

$$L = W\Sigma V^*,$$

- For each $1 \leq k \leq m$, let $e_k = V(\epsilon_k)$
- For each $1 \leq j \leq n$, let $f_j = W(\epsilon_j)$
- Then for any $x = e_1 x^1 + \cdots + e_m x^m \in X$,

$$L(x) = L(e_1 x^1 + \cdots + e_m x^m) = f_1 s_1 x^1 + \cdots + f_r s_r x^r$$

- Therefore, for any $y = f_1 y^1 + \cdots + f_n y^n \in Y$,

$$L^+(y) = L^+(f_1 y^1 + \cdots + f_n y^n) = e_1 s_1^{-1} y^1 + \cdots + e_r s_r^{-1} y^r$$

- In other words,

$$L^+ = W\Sigma^+ V^*$$

Solving a Linear System with Errors

Moore-Penrose Quasi-Inverse

- Let $L : X \to Y$ be a linear map between inner product spaces
- Suppose that, given $y \in Y$, we want to solve

$$L(x) = y,$$

for x but the exact value of y is not known

- If the measured value of y is $y + \Delta y$ and

$$x + \Delta x = L^{-1}(y + \Delta y),$$

then

$$\Delta x = L^{-1}(\Delta y)$$

- The relative error of x can ye estimated in terms of the relative error of y:

$$\frac{|\Delta x|}{|x|} = \frac{|L^{-1}(\Delta y)|}{|y|}\frac{|y|}{|x|} = \frac{|L^{-1}(\Delta y)|}{|y|}\frac{|L(x)|}{|x|} \leq \|L^{-1}\|\|L\|\frac{|\Delta y|}{|y|}$$

Condition Number of Linear Map

- $\|L^{-1}\|\|L\|$ is the **condition number** of the linear map
- It shows how sensitive the error in x is to the error in y
- A linear map is **ill-conditioned** if the condition number is large
- The condition number can be changed by changing the inner product