

Lecture 5

Linear Systems of Equations, Part 2

Virginia Tech Department of Computer Science
Fall 2026 • Sumeet Khatri

Gaussian elimination recap

Reversible row operations act on the entire augmented matrix.

Forward elimination

$$m_{ik} = \frac{a_{ik}}{a_{kk}}, \quad R_i \leftarrow R_i - m_{ik}R_k$$

Remove entries below each pivot.

Back substitution

$$x_i = \frac{b_i - \sum_{j=i+1}^n a_{ij}x_j}{a_{ii}}$$

Recover unknowns from the last row upward.

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 8 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & 0 & -1 & 1 \end{array} \right) \implies \mathbf{x} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

Pivots, consistency, and uniqueness

Read the reduced system

An impossible row $0 = c$, with $c \neq 0$, means no solution.

If consistent: every unknown has a pivot, or free variables give infinitely many solutions.

Choose pivots carefully

Row exchanges avoid zero pivots.

Tiny pivots can amplify roundoff. Scaling and reordering can help.

Outline

Matrix inverses and rank

Singular value decomposition and pseudo-inverse

Numerical rank and truncated SVD

Error analysis

Inverse of a matrix

For $A \in \mathbb{R}^{n \times n}$, an inverse is a matrix $A^{-1} \in \mathbb{R}^{n \times n}$ satisfying

$$A^{-1}A = AA^{-1} = I_n$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad ad - bc \neq 0$$

An inverse exists: **invertible** or **nonsingular**. No inverse exists: **singular**.

Solution of an invertible system

Inverses are unique: If B and C both satisfy the inverse identities for the same square matrix A , then $B = C$.

Theorem: If $A \in \mathbb{R}^{n \times n}$ is invertible, then every $\mathbf{b} \in \mathbb{R}^n$ gives exactly one solution of $A\mathbf{x} = \mathbf{b}$.

$$\mathbf{x} = A^{-1}\mathbf{b}$$

Existence and uniqueness are different questions; invertibility guarantees both.

Rank of a matrix

Row rank

Dimension of the span of the rows.

Column rank

Dimension of the span of the columns.

$$\text{rank}(A) = \dim \text{Row}(A) = \dim \text{Col}(A) \leq \min(m, n)$$

Full row rank: $\text{rank}(A) = m$. Full column rank: $\text{rank}(A) = n$.

Full rank: $\text{rank}(A) = \min(m, n)$.

For a square matrix: invertible $\iff$ full rank

Example: solution via inverse

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

$$AA^{-1} = A^{-1}A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{pmatrix} 5 - 3 \\ -5 + 6 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Inversion using row operations: Gauss–Jordan inversion

Write the unknown inverse as X . Then we want $AX = I_n$.

If we set x_j to be the columns of X , then we require

$$Ax_j = e_j, \quad j = 1, \dots, n$$

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

This is a linear system $\Rightarrow$ use Gaussian elimination!

Theorem 2: Elementary row operations transform $(A \mid I_n)$ into $(I_n \mid X)$ if and only if $A \in \mathbb{R}^{n \times n}$ is invertible. Then $X = A^{-1}$.

$$(A \mid I_n) \longrightarrow (I_n \mid A^{-1})$$

Difference from before: now the right-hand partition is a matrix, not a vector as before!

Example: row reduction to get the inverse

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 1 \end{array} \right) \xrightarrow{R_2 \leftarrow R_2 - \frac{5}{2}R_1} \left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & \frac{1}{2} & -\frac{5}{2} & 1 \end{array} \right)$$

$$\xrightarrow{R_2 \leftarrow 2R_2} \left(\begin{array}{cc|cc} 2 & 1 & 1 & 0 \\ 0 & 1 & -5 & 2 \end{array} \right)$$

Example: row reduction to get the inverse

$$\begin{pmatrix} 2 & 1 & | & 1 & 0 \\ 0 & 1 & | & -5 & 2 \end{pmatrix} \xrightarrow[\begin{matrix} R_1 \leftarrow R_1 - R_2 \\ R_1 \leftarrow \frac{1}{2}R_1 \end{matrix}]{\begin{matrix} R_1 \leftarrow R_1 - R_2 \\ R_1 \leftarrow \frac{1}{2}R_1 \end{matrix}} \begin{pmatrix} 1 & 0 & | & 3 & -1 \\ 0 & 1 & | & -5 & 2 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}, \quad \mathbf{x} = A^{-1}\mathbf{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Rectangular and singular systems

$$A \in \mathbb{R}^{m \times n}, \quad \mathbf{b} \in \mathbb{R}^m, \quad \mathbf{x} \in \mathbb{R}^n$$

A two-sided inverse is unavailable when $m \neq n$. A square matrix may also have dependent columns.

Does an exact solution exist?

If it exists, is it unique?

If exact equality is impossible, what should “solving” mean?

The pseudo-inverse provides one rule for exact solutions, best fits, and a shortest-input choice.

Recall: Eigenvalues and eigenvectors of symmetric matrices

A real symmetric matrix $A \in \mathbb{R}^{n \times n}$ has real eigenvalues and an orthonormal eigenvector matrix Q , with $Q^T Q = I_n$.

$$A\mathbf{q}_i = \lambda_i\mathbf{q}_i, \quad A = Q\Lambda Q^T, \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$$

Use eigenvector coordinates $\mathbf{y} = Q^T \mathbf{x}$ and $\mathbf{c} = Q^T \mathbf{b}$. The system becomes $\Lambda \mathbf{y} = \mathbf{c}$.

If every eigenvalue is nonzero, then we can divide:

$$A^{-1} = Q\Lambda^{-1}Q^T, \quad \Lambda^{-1} = \text{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_n^{-1})$$

$$\mathbf{x} = Q\Lambda^{-1}Q^T \mathbf{b} = \sum_{i=1}^n \frac{\mathbf{q}_i^T \mathbf{b}}{\lambda_i} \mathbf{q}_i$$

Compact singular value decomposition

$$A = U_r \Sigma_r V_r^\top = \sum_{j=1}^r \sigma_j \mathbf{u}_j \mathbf{v}_j^\top, \quad r = \text{rank}(A) > 0$$

$$U_r \in \mathbb{R}^{m \times r}, \quad \Sigma_r \in \mathbb{R}^{r \times r}, \quad V_r \in \mathbb{R}^{n \times r}$$

$$\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r), \quad \sigma_1 \geq \dots \geq \sigma_r > 0, \quad U_r^\top U_r = V_r^\top V_r = I_r$$

σ_j : singular values

$\mathbf{u}_j$: left singular vectors

$\mathbf{v}_j$: right singular vectors

SVD existence and subspaces

Theorem: Every real matrix $A \in \mathbb{R}^{m \times n}$ of rank $r > 0$ has the stated compact SVD. For $A = 0$, use empty sums.

$$A\mathbf{v}_j = \sigma_j \mathbf{u}_j, \quad A^\top \mathbf{u}_j = \sigma_j \mathbf{v}_j, \quad j = 1, \dots, r$$

$$\begin{aligned} \text{Col}(A) &= \text{span}\{\mathbf{u}_1, \dots, \mathbf{u}_r\}, \\ \text{Col}(A^\top) &= \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\}. \end{aligned}$$

Moore–Penrose pseudo-inverse

$$A^+ = V_r \Sigma_r^{-1} U_r^T = \sum_{j=1}^r \frac{1}{\sigma_j} \mathbf{v}_j \mathbf{u}_j^T$$

$$A \in \mathbb{R}^{m \times n} \implies A^+ \in \mathbb{R}^{n \times m}$$

Reciprocate positive singular values only. For $A = 0$, define $A^+ = 0$ with the transposed dimensions.

For an invertible square matrix, $A^+ = A^{-1}$. The pseudo-inverse extends the ordinary inverse.

What does the pseudo-inverse give?

$$\mathbf{x}_* = A^+ \mathbf{b} = \sum_{j=1}^r \frac{\mathbf{u}_j^\top \mathbf{b}}{\sigma_j} \mathbf{v}_j$$

When the system is inconsistent

$$\min_{\mathbf{x}} \|\mathbf{b} - A\mathbf{x}\|_2$$

When the system is consistent

Minimize $\|\mathbf{x}\|_2$ among all possible solutions.

Inner products and vector norms

For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, define the Euclidean inner product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y} = \sum_{j=1}^n x_j y_j$$

Orthogonal means $\mathbf{x}^\top \mathbf{y} = 0$. Orthonormal vectors have unit Euclidean length and are mutually orthogonal.

$$\|\mathbf{x}\|_1 = \sum_{j=1}^n |x_j|, \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^n |x_j|^2}, \quad \|\mathbf{x}\|_\infty = \max_{1 \leq j \leq n} |x_j|$$

Norms are positive except at zero, scale by absolute magnitude, and obey the triangle inequality.

Full-rank pseudo-inverse formulas

Theorem: Rank assumptions can be used to determine formulas for the pseudo-inverse.

Full column rank

$$\text{rank}(A) = n, \quad m \geq n$$

$$A^+ = (A^T A)^{-1} A^T,$$

$$A^+ A = I_n.$$

Full row rank

$$\text{rank}(A) = m, \quad m \leq n$$

$$A^+ = A^T (A A^T)^{-1},$$

$$A A^+ = I_m.$$

Example: a one-row system

$$A = (1 \quad 1), \quad AA^T = (2), \quad A^+ = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$x_1 + x_2 = 2, \quad \mathbf{x}_* = A^+(2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$A^+A = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad AA^+ = I_1$$

Many inputs have the required sum. The pseudo-inverse selects the shortest one.

Example: an inconsistent system

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

An exact fit would require $\alpha = 1$ and $\beta = 1$, but then the last predicted value is 3, not 2. The system is inconsistent.

$$A^T A = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}, \quad A^T \mathbf{b} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}, \quad A^T A \mathbf{x} = A^T \mathbf{b}$$

Example: an inconsistent system

$$(A^T A)^{-1} = \frac{1}{6} \begin{pmatrix} 5 & -3 \\ -3 & 3 \end{pmatrix}, \quad \mathbf{x}_* = \begin{pmatrix} 7/6 \\ 1/2 \end{pmatrix}$$

$$A\mathbf{x}_* = \begin{pmatrix} 7/6 \\ 5/3 \\ 13/6 \end{pmatrix}, \quad \mathbf{r}_* = \mathbf{b} - A\mathbf{x}_* = \begin{pmatrix} -1/6 \\ 1/3 \\ -1/6 \end{pmatrix}$$

$$A^T \mathbf{r}_* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \|\mathbf{r}_*\|_2^2 = \frac{1}{6}$$

Example: an underdetermined system

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$x_1 + x_3 = 1, \quad x_2 + x_3 = 1$$

$$x_3 = t \quad \implies \quad \mathbf{x}(t) = \begin{pmatrix} 1 - t \\ 1 - t \\ t \end{pmatrix}, \quad t \in \mathbb{R}$$

Setting $t = 0$ gives $(1,1,0)^\top$. This is an exact solution, but is it the shortest?

Example: an underdetermined system

$$\|\mathbf{x}(t)\|_2^2 = 2(1-t)^2 + t^2 = 3 \left(t - \frac{2}{3} \right)^2 + \frac{2}{3}$$

Use the full-row-rank formula

$$AA^T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$(AA^T)^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Minimum-norm solution

$$A^+ = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{x}_* = \begin{pmatrix} 1/3 \\ 1/3 \\ 2/3 \end{pmatrix}$$

$t = \frac{2}{3}$ gives squared norm $\frac{2}{3}$, compared with 2 at $t = 0$.

$\text{Null}(A) = \text{span}\{(-1, -1, 1)^T\}$ and $-\frac{1}{3} - \frac{1}{3} + \frac{2}{3} = 0$: the selected solution has no null-space component.

Example: SVD of a singular matrix

The second row is twice the first, so $\text{rank}(A) = 1$.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad \mathbf{u}_1 = \mathbf{v}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \sigma_1 = 5$$

$$A = 5\mathbf{u}_1\mathbf{v}_1^T, \quad A^+ = \frac{1}{5}\mathbf{v}_1\mathbf{u}_1^T = \frac{1}{25} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Neither $A^T A$ nor AA^T is invertible. Use the SVD formula; the special full-rank formulas do not apply.

Example: SVD of a singular matrix

Consistent data

$$\mathbf{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \quad \mathbf{x}_* = \begin{pmatrix} 3/5 \\ 6/5 \end{pmatrix}$$

$$A\mathbf{x}_* = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \quad \mathbf{r}_* = \mathbf{0}$$

Every best fit is exact.

Inconsistent data

$$\mathbf{b} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \quad \mathbf{x}_* = \frac{1}{25} \begin{pmatrix} 13 \\ 26 \end{pmatrix}$$

$$A\mathbf{x}_* = \begin{pmatrix} 13/5 \\ 26/5 \end{pmatrix}, \quad \mathbf{r}_* = \begin{pmatrix} 2/5 \\ -1/5 \end{pmatrix}$$

Elimination gives $0 = -1$; the minimum squared residual is $\frac{1}{5}$.

Sensitivity to changes in the data

Theorem: Fix $A \in \mathbb{R}^{m \times n}$ with rank $r > 0$ and smallest positive singular value σ_r .

$$\mathbf{x}_* = A^+ \mathbf{b}, \quad \tilde{\mathbf{x}}_* = A^+ (\mathbf{b} + \delta \mathbf{b}), \quad \delta \mathbf{x} = \tilde{\mathbf{x}}_* - \mathbf{x}_*$$

$$\delta \mathbf{x} = A^+ \delta \mathbf{b} = \sum_{j=1}^r \frac{\mathbf{u}_j^\top \delta \mathbf{b}}{\sigma_j} \mathbf{v}_j, \quad \|\delta \mathbf{x}\|_2 \leq \frac{\|\delta \mathbf{b}\|_2}{\sigma_r}$$

Numerical rank and singular-value thresholds

$\text{diag}(1, \varepsilon)$ has rank 2 for $\varepsilon \neq 0$ and rank 1 at $\varepsilon = 0$. Its second inverse entry changes from $1/\varepsilon$ to a zero pseudo-inverse entry.

$$J = \{j \in \{1, \dots, r\} : \sigma_j > \tau\}, \quad r_\tau = |J|$$

$$A_\tau = \sum_{j \in J} \sigma_j \mathbf{u}_j \mathbf{v}_j^\top, \quad \mathbf{x}_\tau = A_\tau^+ \mathbf{b} = \sum_{j \in J} \frac{\mathbf{u}_j^\top \mathbf{b}}{\sigma_j} \mathbf{v}_j$$

A common roundoff threshold is $\tau = \max(m, n) \epsilon_{\text{mach}} \sigma_1$. Measurement or model error may call for a different cutoff.

Example: a small positive singular value

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 10^{-8} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 10^{-8} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (\text{Exact solution})$$

Now introduce a cut-off:

$$\tau = 10^{-6} : \quad \mathbf{x}_\tau = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{b} - A\mathbf{x}_\tau = \begin{pmatrix} 0 \\ 10^{-8} \end{pmatrix}$$

Residual, forward error, and backward error

Let $Ax = b$ be nonsingular and let $\hat{x}$ be a computed approximation.

Residual

$$r = b - A\hat{x}$$

Computable from the original data.

Forward error

$$\hat{x} - x$$

Usually unknown because the exact solution is unknown.

$$A\hat{x} = b - r$$

With A fixed and $b \neq 0$, $\|r\|/\|b\|$ is a relative data backward error.

Scaled residual and backward error

Theorem: Let $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\hat{\mathbf{x}} \in \mathbb{R}^n$. Use the spectral and Euclidean norms.

$$d = \|A\|_2 \|\hat{\mathbf{x}}\|_2 + \|\mathbf{b}\|_2, \quad \eta = \frac{\|\mathbf{b} - A\hat{\mathbf{x}}\|_2}{d}$$

If $d = 0$, the residual is zero and we define $\eta = 0$.

If $d > 0$, then η is the smallest $\epsilon \geq 0$ for which perturbations can satisfy all three conditions:

$$(A + \Delta A)\hat{\mathbf{x}} = \mathbf{b} + \Delta \mathbf{b}, \\ \|\Delta A\|_2 \leq \epsilon \|A\|_2, \quad \|\Delta \mathbf{b}\|_2 \leq \epsilon \|\mathbf{b}\|_2.$$

Induced matrix norms

For $A \in \mathbb{R}^{m \times n}$ and $p \in \{1, 2, \infty\}$, use the same norm type for inputs and outputs.

$$\|A\|_p = \max_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p$$

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \quad \|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$$

$$\|A\|_2 = \max_j \sigma_j \quad (\text{largest singular value, aka spectral norm})$$

$$\|A\mathbf{x}\|_p \leq \|A\|_p \|\mathbf{x}\|_p, \quad \|AB\|_p \leq \|A\|_p \|B\|_p$$

Relative error bound

Theorem: Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and $\mathbf{b} \neq 0$. For any approximation $\hat{\mathbf{x}} \in \mathbb{R}^n$, define

$$\mathbf{x} = A^{-1}\mathbf{b}, \quad \mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}}$$

For every vector norm and its induced matrix norm, the exact residual satisfies

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa(A) \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$

Interpret a small residual together with the condition number.

Condition number

For nonsingular $A \in \mathbb{R}^{n \times n}$ and an induced matrix norm:

$$\kappa(A) = \|A\| \|A^{-1}\|$$

$$\kappa_2(A) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)} \quad (\text{spectral norm})$$

A large condition number permits large relative solution changes from small relative changes in the data.

$$\kappa(A) \geq 1, \quad \kappa(cA) = \kappa(A) \quad (c \neq 0)$$

Condition estimation does not require forming an explicit inverse.

Example: condition number

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 10^{-8} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 10^{-8} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \hat{\mathbf{x}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\mathbf{r} = \begin{pmatrix} 0 \\ 10^{-8} \end{pmatrix}, \quad \frac{\|\mathbf{r}\|_2}{\|\mathbf{b}\|_2} = \frac{10^{-8}}{\sqrt{1 + 10^{-16}}} \approx 10^{-8}$$

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{1}{\sqrt{2}}, \quad \kappa_2(A) = 10^8$$

The small coefficient makes a large error in x_2 almost invisible in the residual.

Example: condition number

Use the same A , $\mathbf{b}$, $\mathbf{x}$, and $\hat{\mathbf{x}}$ as Example 10.

$$\|A\|_{\infty} = 1, \quad \|A^{-1}\|_{\infty} = 10^8, \quad \kappa_{\infty}(A) = 10^8$$

$$\frac{\|\mathbf{r}\|_{\infty}}{\|\mathbf{b}\|_{\infty}} = 10^{-8}, \quad \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_{\infty}}{\|\mathbf{x}\|_{\infty}} = 1$$

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_{\infty}}{\|\mathbf{x}\|_{\infty}} = 1 = 10^8 \cdot 10^{-8} = \kappa_{\infty}(A) \frac{\|\mathbf{r}\|_{\infty}}{\|\mathbf{b}\|_{\infty}}$$

Simultaneous matrix and data perturbations

Theorem: Let $A \in \mathbb{R}^{n \times n}$ be nonsingular, $\mathbf{b} \neq 0$, and $A\mathbf{x} = \mathbf{b}$. Use a vector norm and its induced matrix norm.

$$\epsilon_A = \frac{\|\Delta A\|}{\|A\|}, \quad \epsilon_b = \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|}$$

If $\kappa(A)\epsilon_A < 1$, then $A + \Delta A$ is nonsingular. Its unique perturbed solution satisfies

$$(A + \Delta A)(\mathbf{x} + \Delta \mathbf{x}) = \mathbf{b} + \Delta \mathbf{b}$$

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\kappa(A)}{1 - \kappa(A)\epsilon_A}(\epsilon_A + \epsilon_b)$$

The guarantee weakens as the denominator approaches zero.