Solving systems of linear equations

This is one of the most common tasks, and care must be taken to do it well and efficiently.

Applications (1) Newton's method

$$J(x_n)(x_{n+1} - x_n) = -F(x_n)$$

$F : \mathbb{R}^n \to \mathbb{R}^n$, $J : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ (the space of real $n\times n$ matrices)

(2) Data analysis - least squares

(3) Optimization

(4) BVPs in PDE

Given $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, find $x \in \mathbb{R}^n$

s.t. $\quad Ax = b$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{kk} & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

One obvious approach: Find the <u>inverse</u> of A, $A^{-1} \in \mathbb{R}^{n\times n}$, satisfying

$$A^{-1}A = AA^{-1} = I \qquad I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

The inverse matrix A^{-1} exists iff $\det(A) \neq 0$, in which case we call A <u>nonsingular</u> or <u>invertible</u>.

If A is nonsingular, then

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

Is this the way to do it? <u>No</u>. Finding A^{-1} directly is very expensive (time-consuming)

<u>Constructing A^{-1}</u>. Say you have a method for solving $Ax = b$. Then solve

$$\left.\begin{array}{l} A x^1 = e^1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ A x^2 = e^2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ \quad \vdots \qquad \vdots \\ A x^n = e^n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \end{array}\right\} \text{ and } A^{-1} = [x^1, \dots, x^n]$$

unit vectors

(matrix of column vectors)

If nothing special is done, this is $n \times$ the cost of directly solving $Ax = b$ in the 1st case. [You might want to do this if you will use A^{-1} repeatedly]

<u>or</u> Cramer's Rule $\quad x_i = \dfrac{\det(A_i^b)}{\det(A)}$

$$A_i^b = [A^1, \dots, A^{i-1}, b, A^{i+1}, \dots, A^n]$$

replace i^{th} column.

$\det(A)$ costs $\sim n!$ floating point opns (flops) $(+, -, \times, \div)$

Much, much cheaper ways.

Gaussian elimination (C.F. Gauss)

Successive elimination of rows. $Ax = b$

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 16 \\ -3 \end{bmatrix}; \quad \begin{bmatrix} r_1^0 \\ r_2^0 \\ r_3^0 \end{bmatrix} = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 2 & 4 & 2 & 16 \\ -1 & 5 & -4 & -3 \end{array}\right]$$

↑ add multiples of 1^{st} row to eliminate the 1^{st} elements of lower rows.

i.e.

$$r_2^1 \leftarrow r_2^0 - 2 \cdot r_1^0 \quad (1)$$

$$r_3^1 \leftarrow r_3^1 + 1 \cdot r_1^0 \quad (2)$$

or

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 2 & 0 & 4 \\ 0 & 6 & -3 & 3 \end{array}\right]$$

This turns out to be equivalent to successive multiplication by lower triangular matrices L_1 & L_2.

$$L = \begin{pmatrix} l_{11} & 0 & & 0 \\ l_{21} & l_{22} & 0 & \\ \vdots & & \ddots & 0 \\ l_{n1} & \cdots & & l_{nn} \end{pmatrix} = \begin{bmatrix} \diagdown & 0 \\ & \end{bmatrix}$$

$$L_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I + \mu_{21} \underbrace{e^2 (e^1)^T}_{A^{21}}, \quad \mu_{21} = -2$$

A^{21}, A^{31} ← rank-one matrices. $(ab^T)_{ij} = a_i b_j$

$$L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = I + \mu_{31} \underbrace{e^3 (e^1)^T}_{A^{31}}$$

$\mu_{31} = 1$

$$e^k (e^l)^T = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k^{th} \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & & 0 \\ & 1_{k,l} & \\ 0 & & 0 \end{pmatrix}$$

(↑ l^{th})

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 2 & 0 & 4 \\ 0 & 6 & -3 & 3 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & -3 & -9 \end{array}\right]; \quad L_3 = I - 3 e^3 (e^2)^T = I + \mu_{32} A^{32}$$

We now have

$$Ux = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ -9 \end{bmatrix} = \tilde{b}$$

Solve by "back-substitution"

upper triangular

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Before moving on, we had

$$Ax = b$$

$$L_3 L_2 L_1 A x = Ux = L_3 L_2 L_1 b = \tilde{b}$$

$$L_3 L_2 L_1 A = U \underset{\text{if } L_i^{-1} \text{ exists}}{\Longrightarrow} A = L_1^{-1} L_2^{-1} L_3^{-1} U$$

<u>Fact</u> $\left(I + \mu e^k (e^l)^T\right)^{-1} = I - \mu e^k (e^l)^T$

$k \neq l$

<u>Check</u> $\left(I + \mu e^k (e^l)^T\right)\left(I - \mu e^k (e^l)^T\right)$

$$= I + (\mu - \mu) e^k (e^l)^T - \mu^2 e^k \underbrace{(e^l)^T e^k}(e^l)^T$$

$$= I$$

[special case of the Sherman-Morrison Formula]

$e^l \cdot e^k = \delta_{kl}$

$$L_1^{-1} L_2^{-1} L_3^{-1} = \left(I - \mu_1 e^2 (e^1)^T\right)\left(I - \mu_2 e^3 (e^1)^T\right)\left(I - \mu_3 e^3 (e^2)^T\right)$$

$$= I + \alpha_1 e^2 (e^1)^T + \alpha_2 e^3 (e^1)^T + \alpha_3 e^3 (e^2)^T$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ \alpha_1 & 1 & 0 \\ \alpha_2 & \alpha_3 & 1 \end{bmatrix}$$

"lower unit triangular"

$A = LU$ "matrix decomposition"

Triangular Matrices

$L \in \mathbb{R}^{n \times n}$ is called "lower triangular" if

$$l_{ij} = 0 \quad \forall \; 1 \le i < j \le n$$

i.e. $L = \begin{bmatrix} l_{11} & 0 & 0 \\ \vdots & l_{uu} & 0 \\ l_{in} & \cdots & l_{nn} \end{bmatrix}$

"unit lower triangular" if

$$l_{11} = l_{22} = \cdots = l_{uu} = 1$$

& similarly for upper triangular matrices.

Properties of lower triangular matrices

(1) If L_1, L_2 (unit) lower triangular matrices then $L_3 = L_1 L_2$ is a (unit) lower Δ matrix.

(2) lower triangular matrices are n.s. iff

$l_{ii} \neq 0 \quad 0 \le i \le n$. why?

$$\det L = \prod_{i=1}^{n} l_{ii}$$

(3) If L, a (unit) lower Δ matrix, is n.s. then L^{-1} is also a (unit) lower Δ matrix

Pf. By induction on matrix dimension n:

$\underline{n=2}$ $\quad L = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$ $\quad a, c \neq 0$ since L is n.s.

By direct calculation $L^{-1} = \frac{1}{ac}\begin{pmatrix} c & 0 \\ -b & a \end{pmatrix}$ ✓

Assume true for any $n \times n$ lower Δ matrix

Let $L \in \mathbb{R}^{(n+1)\times(n+1)}$. Write L as

$$L = \left[\begin{array}{c|c} L_1 \in \mathbb{R}^{n\times n} \text{ lower } \Delta & 0 \\ \hline - r^T - & \alpha \end{array}\right] ; \quad L^{-1} = \left[\begin{array}{c|c} L_1' \in \mathbb{R}^{n\times n} & c \\ \hline - s^T - & \mu \end{array}\right]$$

$$LL^{-1} = I \in \mathbb{R}^{(n+1)\times(n+1)} = \left[\begin{array}{c|c} I \in \mathbb{R}^{n\times n} & 0 \\ \hline - 0 - & 1 \end{array}\right]$$

$$= \left[\begin{array}{c|c} L_1 L_1' & L_1 c \\ \hline r^T L_1' + \alpha s^T & \beta \end{array}\right]$$

By assumption $L_1 L_1' = I$

$\Rightarrow L_1' = L_1^{-1}$ is lower Δ

$L_1 c = 0$, L_1 n.s. $\Rightarrow c \equiv 0$ ✓ done.

<u>Also</u> $r^T L_1' + \alpha s^T = 0$