Lecture 16: Numerical Linear Algebra

Outline

0) Newton Demo for Linear systems:
1) Overview of Linear Algebra
2) Basics: Vector and Matrix Norms
3) The condition number of a matrix cond(A)
4) The condition number and error estimation in A$\underline{x}$=$\underline{b}$

Newton's method for $\underline{F}(\underline{x})=\underline{0}$

Example: Linear problem

$\underline{F}(\underline{x})=\underline{0}$ $\quad \underline{F}(\underline{x})=A\underline{x}-\underline{b}$ $\quad F(\underline{x}^*)=A\underline{x}^*-\underline{b}=\underline{r}$

$F_1(\underline{x}) = ax_1 + bx_2 - b_1$

$F_2(\underline{x}) = cx_1 + dx_2 - b_2$

$J_{ij} = \frac{\partial F_i}{\partial x_j}$ $\quad J = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = A$

guess $\underline{x} = \underline{x}^*$

$\underline{x}_{n+1} = \underline{x}^* - J^{-1}(x^*)\,\underline{F}(x^*)$

$\underline{x}_{n+1} = x^* - A^{-1}(Ax^* - b) =$

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

Overview of Linear Algebra

subject:	Linear Systems	Least Squares	Eigen Problems
Equations:	$A\underline{x}=\underline{b}$	$A^TA\underline{x}=A^T\underline{b}$	$A\underline{x}=\lambda\underline{x}$
Algorithms:	Elimination (Gauss,GJordan)	Gram-Schmidt Householder	factor $P(\lambda)=\|A-\lambda I\|$ Find $N(A-\lambda_i I)$
Factorizations:	$PA=LU$ $PA=LDL^T$ $A=CC^T$	$A = QR$ $Q^TQ=I$	$AS = S\Lambda$ $A = S\Lambda S^{-1}$ $A = Q\Lambda Q^T$ $A = U\Sigma V^T$

Outline of Lectures for Numerical Linear Algebra

1) Numerical issues and error analysis for NLA (this lecture)

2) Direct methods for solving square linear systems $A\underline{x}=\underline{b}$
   Gaussian Elimination with Partial Pivoting and the LU factorization

3) Direct methods for solving linear least squares problems
   Normal Equations, Orthogonalization and the QR factorization

4) Iterative methods for solving $A\underline{x}=\underline{b}$ and relationship to eigenvalue problems
   Simple iterative schemes
   Power method and variations for eigenvalues/eigenvectors
   Introduction to Krylov subspace schemes

Numerical Issues:

Floating Point Errors:

Considerable amount of mults & adds can compound FP errors

Need Stable schemes that don't amplify FP Errors

Ill-Conditioning: Ill condition

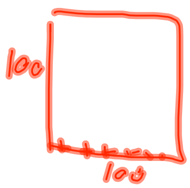
Possibility of "near-singular" matrices (badly behaved, particularly with FP errors)

$A\underline{x}=\underline{b}$ $\begin{bmatrix} | & | \\ a_1 & a_2 \\ | & | \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\underline{a}_1 + x_2\underline{a}_2$

Performance and scaling:

Solution of large systems can be extremely expensive computationally

e.g. $O(N^3)$ -- need fast algorithms for big problems



A is invertible x is unique

square $n\times n$ A is singular

no solution $b \notin C(A)$

∞# solution $b \in C(A)$

$\underline{x} = \underline{x}_p + N_s$

Quantifying Errors and *Condition Number* of a matrix:

Motivation: Need a scalar measure of how "singular" a matrix is

Why not the Determinant |A|?

if A is singular $|A| = 0$

$I = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$ $|I| = 1$

$|\varepsilon I| = \varepsilon^3$

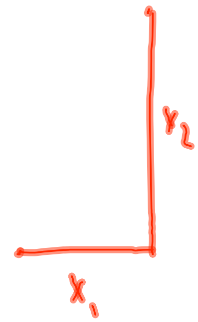
$I^{10{,}000 \times 10000s}$

$\varepsilon I = \begin{bmatrix} \varepsilon & & \\ & \varepsilon & \\ & & \varepsilon \end{bmatrix}$ $|\varepsilon I| = \varepsilon^n$

Quantifying Errors and *Condition Number* of a matrix: Vector Norms

Definition: the p-norm of a vector is defined as $||\underline{x}||_p$

$$||\underline{x}||_p = \left[\sum_i |x_i|^p\right]^{1/p}$$



Example: x=[1 -1 1]'

1-norm: (Manhattan Distance) $= \sum |x_i| = 3$

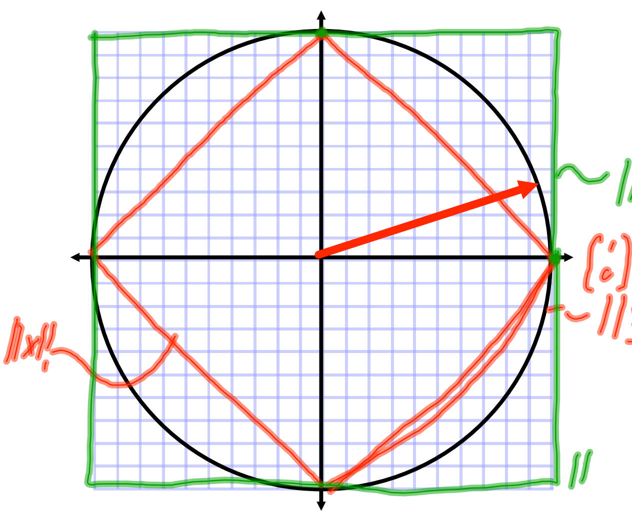
$p=2$

2-norm: (Euclidean distance) $||x||_2 = \sqrt{\underline{x}^T\underline{x}} = \sqrt{3}$

∞-norm: (max(abs(x))) $||x||_\infty = 1$

Quantifying Errors and *Condition Number* of a matrix: Vector Norms

The unit "sphere" in the 1,2 and inf-norm in R^2



in 1-norm $|x_1|+|x_2|=1$

∞-norm $\begin{bmatrix}\alpha\\1\end{bmatrix} \to \begin{bmatrix}1\\\alpha\end{bmatrix}$

$\underline{x} = \begin{bmatrix}1\\1\end{bmatrix}$

$||\underline{x}||_1 = 2 \to N$

$||\underline{x}|| = \sqrt{2}$ $\quad \sqrt{N}$

$||\underline{x}|| = 1$ $\quad 1$

$||x||_1 \geq ||x||_2 \geq ||\underline{x}||_\infty$

Properties of Vector Norms

All vector norms have 3 fundamental properties

$\|x\| > 0$ if $x \neq 0$

$\alpha \in \mathbb{R}$ $\|\alpha x\| = |\alpha| \, \|x\|$

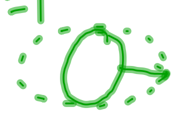
$\|x + y\| \le \|x\| + \|y\|$

Matrix p-norms

"Induced" by vector p-norms

Definition: the p-norm of a matrix $\|A\|_p$

$A = \begin{bmatrix} 2 & \\ & 1 \end{bmatrix}$ $\|A\|_2 = 2$ $\quad \|A\|_p = \max_{x} \dfrac{\|Ax\|_p}{\|x\|_p}$ $\quad \|x\|_p = 1$



Point: describes the maximum distortion of the "unit-sphere"

Quick Definitions (and matlab): $\begin{bmatrix} 1 & 1 \\ a_1 & a_2 \\ 1 & 1 \end{bmatrix}$ $\|A\|_\infty = \|A^T\|$

$\|A\|_1$ = norm(A,1) = max(sum(abs(A))) (is max 1-norm of columns of A)

$\|A\|_2$ = norm(A) = max(svd(A)) (maximum singular value of A)

$\|A\|_\infty$ = norm(A,'inf') = max(sum(abs(A')) (is max 1-norm of rows of A or $\|A^T\|_1$

Matrix p-norms

Examples:

1) $||I|| =$ $\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$ $||I||_1 = 1 = ||I||_\infty = ||I||_2$

2) $||D|| =$ $\begin{bmatrix} 2 & & \\ & 1 & \\ & & -3 \end{bmatrix}$ $||D||_1 = 3 = ||D||_\infty = ||D||_2 = 3$

$\begin{bmatrix} d_1 & & \\ & d_2 & \\ & & d_3 \end{bmatrix}$ $||D|| = \max(\text{abs}(\text{diag}(D)))$

3) let Q by orthonormal s.t. $Q^TQ=I$, then $||Q||_2=$ $\underline{y} = Q\underline{x}$ $||\underline{y}||_2 = ||Q\underline{x}||_2$

$y^Ty = x^Tx$

$||y||_2 = ||\underline{x}||_2$

$||Q\underline{x}||_2^2 = (Qx)^T(Qx)$

$||Q||_2 = \max \dfrac{||Qx||_2}{||x||_2} = 1$

$\underline{x}^T \underbrace{Q^TQ} \underline{x} = \underline{x}^T\underline{x} = ||x||$

4) A =

Properties of Matrix Norms

All Matrix norms are defined by 3 fundamental properties

$||A|| > 0$ for $A \neq 0$

$||\alpha A|| = |\alpha| \, ||A||$

$||A+B|| \leq ||A|| + ||B||$

In addition, matrix p-norms satisfy

1 product rule

$||AB|| \leq ||A|| \, ||B||$

$||A\underline{x}|| \leq ||A|| \, ||x||$

Matrix p-norms

Proof: $||A||_1$ = max(sum(abs(A))) (is max 1-norm of the columns of A)

$$A\underline{x} = \begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \underline{a}_1 + x_2 \underline{a}_2 + x_3 \underline{a}_3$$

$||x||_1 = 1$

sum $|\,|\,|\,|$ . . . . $||a||_1$

$$\max ||A\underline{x}|| = ||x_1\underline{a}_1 + x_2\underline{a}_2 + x_3\underline{a}_3||_1 \le ||x_1\underline{a}_1||_1 + ||x_2\underline{a}_2||_1 + ||x_3 a_3||$$

$$\le |x_1|\,||a_1|| + |x_2|\,||a_2|| + |x_3|\,||a_3||$$

$||a_i||_{max}$

$$\le |x_1|\,||a||_{max} + |x_2|\,||a||_{max} \ldots$$

$$= ||a||_{max} \sum_i |x_i| \qquad = 1$$

$$= ||a||_{1,max}$$

Matrix p-norms

Proof: $||A||_2$ = max(singular value)

$$A = U\Sigma V^T \qquad AV = U\Sigma$$

$V^TV = I$

$U^TU = I$

Σ = diagonal

$$||A\underline{v}|| = ||U\Sigma||$$

$$||A||\,||V||_2 = ||U||_2\,||\Sigma||$$

$\qquad 1 \qquad\qquad 1$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix}$$

$$||A||_2 \le ||\Sigma||_2 = \max(\sigma_i) = \sigma_1$$

$\sigma_i > 0$

The condition number of a matrix A

Definition: the condition number of a matrix A , cond(A) (or sometimes κ(A)) is $||A||\ ||A^{-1}||$

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

$A\underline{x} = \underline{b}$

Condition number of special matrices

1) cond(I) = $||I||\,||I^{-1}|| = 1$ $\quad cond(A) \geq 1$ $\quad \underline{x} = A^{-1}\underline{b}$

2) cond(D) = $||D||\,||D^{-1}|| = \frac{D_{max}}{D_{min}}$ $\quad D = \begin{bmatrix} 1 & & \\ & 2 & \\ & & -3 \end{bmatrix}$ $\quad D^{-1} = \begin{bmatrix} 1 & & \\ & 1/2 & \\ & & -1/3 \end{bmatrix}$

3) cond(A) for A singular: $cond(A) = \infty$

4) cond(αA) =

The condition number of a matrix A

Example: A = [(1+α) 1 ; 1 (1+α)] $\varepsilon_{mach} < \alpha << 1$

The condition number and error analysis of A$\underline{x}$=$\underline{b}$

Relative forward and backwards error

$$A\underline{x} = \underline{b}$$

$$\underline{x}^* - \underline{x} = \Delta\underline{x}$$

$$A\Delta x = \Delta\underline{b}$$

$$A\underline{x}^* = \underline{b} + \Delta\underline{b}$$

relative errors $\dfrac{\|\Delta\underline{x}\|}{\|x\|}, \dfrac{\|\Delta b\|}{\|b\|}$

$$\Delta\underline{x} = A^{-1}\Delta\underline{b} \qquad \frac{\|\Delta x\|}{\|\underline{b}\|} \le \frac{\|A^{-1}\|\,\|\Delta b\|}{\|b\|}$$

$$\dot{\underline{b}} = A x$$

$$\therefore \|b\| \le \|A\|\,\|x\|$$

$$\frac{\|\Delta x\|}{\|A\|\,\|x\|} \le \frac{\|A^{-1}\|\,\|\Delta b\|}{\|b\|}$$

$$\frac{\|\Delta x\|}{\|\underline{x}\|} = \text{cond}(A)\frac{\|\Delta b\|}{\|b\|}$$

The condition number and error analysis of A$\underline{x}$=$\underline{b}$

Relative correction error and the residual for iterative methods