Note Title 2/2/2015

Iterative Methods: Produce a sequence x_n of approximations to the solution x^*

Direct Methods: Produce the solution x^* (if we had infinite precision) after a finite number of steps.

Convergence of iterative methods (assuming ∞ precision)

Converges if $\|x_n - x^*\| \longrightarrow 0$

linear convergence if there exists C such that $0 < C < 1$

and $\|x_{n+1} - x^*\| < C\,\|x_n - x^*\|$ for $n \geq N$

Convergence of order β if

$$\|X_{n+1} - X^*\| < \alpha \|X_n - X^*\|^\beta \qquad \text{for} \quad n \geq N$$

Quadratic convergence if $\beta = 2$ $\qquad \alpha > 0$

Sometimes convergence depends on the initial guess x_0

Computational complexity of a direct method = number of operations required to get the solution

Linear systems

$\mathbb{R}$ = set of real numbers

$\mathbb{R}^n$ = set of n-vectors

$x \in \mathbb{R}^n$ then $x = (x_1, \ldots, x_n)$ each $x_i \in \mathbb{R}$ or $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

<u>Norm</u>: A norm on $\mathbb{R}^n$ is a real valued function $\|x\|$ such that:

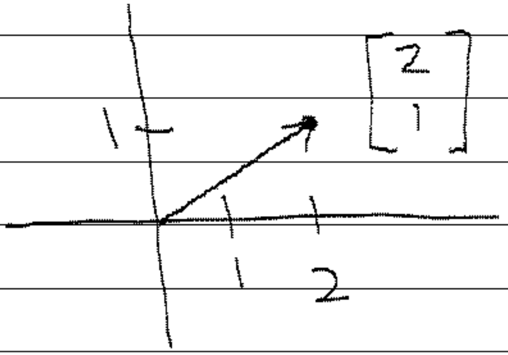
1) $\|x\| \geq 0$ for all $x \in \mathbb{R}^n$ and $\|x\| = 0 \iff x = 0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

2) $\|\lambda x\| = |\lambda| \, \|x\|$ for all $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

3) $\|x+y\| \leq \|x\| + \|y\|$ for all $x, y \in \mathbb{R}^n$

<u>Examples of norms in $\mathbb{R}^n$</u>

1) $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$



Euclidean norm

2) $\|X\|_\infty = \max_{1 \le i \le n} |X_i|$ $\qquad \left\| \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} \right\|_\infty = 4$

3) $\|X\|_1 = \sum_{i=1}^{n} |X_i|$ $\qquad \left\| \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} \right\|_1 = 3+4+1 = 8$

Matrices a k by n matrix A is an object of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix}$$

Ex $A = \begin{bmatrix} 2 & -1 & 0 & 1 \\ 5 & 3 & \boxed{1} & -1 \\ -1 & 1 & 2 & 0 \end{bmatrix}$ is a 3 by 4 matrix or 3×4

Notation: If A is a matrix, a_{ij} is the component of A in the ith row and jth column. In the above example

$a_{23}=1$. Sometimes we use the notation $[A]_{ij}=a_{ij}$

<u>Norms in $\mathbb{R}^{k\times n}$</u>: Same as with vectors, a norm is a real valued function $\|A\|$ that satisfies

1) $\|A\|\geq 0$ for all $A\in\mathbb{R}^{k\times n}$. $\|A\|=0 \Leftrightarrow A=0=\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$
2) $\|\lambda A\|=|\lambda|\,\|A\|$ for all $\lambda\in\mathbb{R}$ and $A\in\mathbb{R}^{k\times n}$
3) $\|A+B\|\leq\|A\|+\|B\|$ for all A and $B\in\mathbb{R}^{k\times n}$

<u>Need to know</u>

1) Addition of vectors
2) Addition of matrices
3) Matrix-vector multiplication
4) Scalar-vector multiplication

5) Scalar matrix multiplication 6) Matrix-matrix multiplication

Vector induced norms on matrices

$A \in \mathbb{R}^{k \times n}$ $\quad x \in \mathbb{R}^n$

$Ax \in \mathbb{R}^k$

$$\begin{bmatrix} 2 & -1 & 0 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix} = \begin{bmatrix} 2(2)+(-1)(-4)+0(1) \\ 3(2)+1(-4)+(-1)1 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \end{bmatrix}$$

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

<u>Example</u>: $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$. What is $\|A\|_2$?

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2$$

$$Ax = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ x_2 \end{bmatrix} \qquad \|Ax\|_2 = \sqrt{4x_1^2 + x_2^2}$$

What $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ with $x_1^2 + x_2^2 = 1$ gives $4x_1^2 + x_2^2$ as large as possible?

$$x_1 = \cos t \qquad x_1^2 + x_2^2 = \cos^2 t + \sin^2 t$$

$$x_2 = \sin t$$

$$f(t) = 4x_1^2 + x_2^2 = 4\cos^2 t + \sin^2 t = 3\cos^2 t + 1 \qquad t = 0$$

$$f(0) = 4 \qquad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Then $\boxed{\|A\|_2} = \max\limits_{\|x\|_2 = 1} \|Ax\|_2 = \left\|A\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\|_2 = \left\|\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right\|_2 = \boxed{\sqrt{2^2 + 0^2} = 2}$

<u>Example</u>: $\|A\|_\infty = \max\limits_{\|x\|_\infty = 1} \|Ax\|_\infty$

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n \end{bmatrix}$$

$\|X\|_\infty$ then $|x_i| \leq 1$ then

(1) $|a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n| \leq |a_{i1}x_1| + |a_{i2}x_2| + \cdots + |a_{in}x_n|$

$\leq |a_{i1}| + |a_{i2}| + \cdots + |a_{in}|$

$$x_j = \begin{cases} 1 & \text{if } a_{ij} \geq 0 \\ -1 & \text{if } a_{ij} < 0 \end{cases}$$

(2) $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = |a_{i1}| + |a_{i2}| + \cdots + |a_{in}|$

If $\|X\|_\infty \leq 1$

$$\|Ax\|_\infty = \max_{1 \leq i \leq n} \left| [Ax]_i \right| \leq \max_{1 \leq i \leq n} |a_{i1}| + |a_{i2}| + \cdots + |a_{in}|$$

$$\|A\|_\infty \le \max_{1 \le i \le n} |a_{i1}| + |a_{i2}| + \cdots + |a_{in}|$$

Let i_0 be such that $|a_{i_0 1}| + \cdots + |a_{i_0 n}| = \max_{1 \le i \le n} |a_{i1}| + \cdots + |a_{in}|$

$$x_j = \begin{cases} 1 & \text{if } a_{i_0 j} \geq 0 \\ -1 & \text{if } a_{i_0 j} < 0 \end{cases}$$ then

$$\|A\|_\infty \geq \|Ax\|_\infty = |a_{i_0 1}| + \cdots + |a_{i_0 n}| = \max_{1 \le i \le n} |a_{i1}| + |a_{i2}| + \cdots + |a_{in}|$$

$$\boxed{\|A\|_\infty = \max_{1 \le i \le n} |a_{i1}| + |a_{i2}| + \cdots + |a_{in}|}$$

Obs: $\|A\| \geq \dfrac{\|Ax\|}{\|x\|}$ $\quad x \in \mathbb{R}^n$, $A \in \mathbb{R}^{k\times n}$

then

$\boxed{\|Ax\| \leq \|A\|\,\|x\|}$ for all $x \in \mathbb{R}^n$ $A \in \mathbb{R}^{k\times n}$

Things we need to know

1) Eigenvalues and eigenvectors
2) Invertible matrices.
3) Inverse of a matrix

Condition number of a matrix

A is invertible $\quad A^{-1}$ is the inverse of A

$AA^{-1} = I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$

$Ax = b$ A & b given
goal find x

$A(x+\delta x) = b + \delta b$ A given b given δb error

How big is the error in the output, i.e. δx?

$Ax + A\delta x = b + \delta b$ but $Ax = b$ then

$A\delta x = \delta b$

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

(a) $\boxed{\|\delta x\| = \|A^{-1}\delta b\| \le \|A^{-1}\| \, \|\delta b\|}$

$b = Ax$ then $\|b\| \le \|A\| \|x\|$ then (b) $\boxed{\dfrac{1}{\|x\|} \le \|A\| \dfrac{1}{\|b\|}}$

$$\boxed{\frac{\|\delta x\|}{\|x\|} \le \|A^{-1}\|\|A\| \frac{\|\delta b\|}{\|b\|}} \quad (*)$$

relative error of output

relative error of input

Def: $K(A) = \|A^{-1}\| \|A\|$ = condition number of the matrix. $(*)$ is a bound. You shoud expect

$$\frac{\|\delta x\|}{\|x\|} \approx K(A) \frac{\|\delta b\|}{\|b\|}$$

The relative error in the input is amplified by a factor of $K(A)$