Iterative Techniques in Matrix Algebra.

We are interested in solving large linear systems $Ax = b$.

Suppose the matrix A has a high percentage of zeros. We would like to take advantage of this sparse structure to reduce the amount of computational work required.

Unfortunately, Gaussian elimination is often unable to take advantage of the sparse structure.

For this reason, we consider iterative techniques.

To estimate how well a particular iterate approximates the true solution we will need some measurement of distance. This motivates the notion of a norm:

Defn A vector norm on $\mathbb{R}^n$ is a function, $\|\cdot\|$, from $\mathbb{R}^n$ into $\mathbb{R}$ with the following properties:

(i) $\|x\| \geq 0$ for all $x \in \mathbb{R}^n$

(ii) $\|x\| = 0 \iff x = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0$

(iii) $\|\alpha x\| = |\alpha| \, \|x\|$ for all $\alpha \in \mathbb{R}$ and $x \in \mathbb{R}^n$

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

Defn The l_2 or Euclidean norm of the vector x is given by

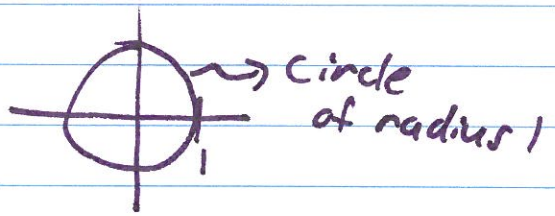
$$\|x\|_2 = \left\{ \sum_{i=1}^{n} x_i^2 \right\}^{1/2}$$

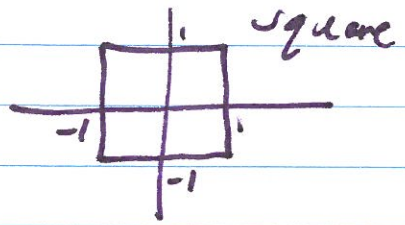
(This represents the usual notion of distance).

Defn The infinity or max norm of a vector x is given by

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|$$

ex Plot $\{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$, $\{x \in \mathbb{R}^2 : \|x\|_\infty = 1\}$





It is straightforward to check that the max norm satisfies the definition of a norm.

To check that the ℓ_2-norm satisfies

$$\|x+y\|_2 \le \|x\|_2 + \|y\|_2$$

requires a famous inequality ...

Cauchy – Schwarz Inequality

For each $x, y \in \mathbb{R}^n$

$$\sum_{i=1}^{n} |x_i y_i| \leq \left\{ \sum_{i=1}^{n} x_i^2 \right\}^{1/2} \left\{ \sum_{i=1}^{n} y_i^2 \right\}^{1/2}$$

Now we can prove that

$$\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$$

Proof: $\|x+y\|_2^2 = \sum_{i=1}^{n} (x_i + y_i)^2$

$$= \sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} y_i^2$$

$$\leq \sum_{i=1}^{n} x_i^2 + 2\|x\|_2 \|y\|_2 + \sum_{i=1}^{n} y_i^2$$

$$= \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2^2$$

$$= \left(\|x\|_2 + \|y\|_2\right)^2$$

$\therefore \|x+y\|_2 \leq \|x\|_2 + \|y\|_2$.

Having defined the idea of worm, we can define the distance between 2 vectors:

Def. If $x, y \in \mathbb{R}^n$ then the

ℓ_2 distance between x and y is defined by

$$\|x-y\|_2 = \left\{ \sum_{i=1}^{n} (x_i - y_i)^2 \right\}^{1/2}$$

and the ℓ_∞ distance between x and y is defined by

$$\|x-y\|_\infty = \max_{1 \le i \le n} |x_i - y_i|$$

Iterative techniques generate a sequence of vectors

A sequence $\{x^{(k)}\}_{k=1}^{\infty}$ of vectors in $\mathbb{R}^n$ is said to converge to x with respect to the norm $\|\cdot\|$ if, given any $\varepsilon > 0$, there exists an integer $N(\varepsilon)$ such that

$$\|x^{(k)} - x\| < \varepsilon \quad \text{for all} \quad k \geq N(\varepsilon).$$

Checking convergence in the max norm is facilitated by the following theorem:

Thm. The sequence of vectors $\{x^{(k)}\}$ converges to x in $\mathbb{R}^n$ with respect to $\|\cdot\|_\infty$ if and only if $\lim_{k\to\infty} x_i^{(k)} = x_i$ for each i.

Proof. Suppose $\{x^{(k)}\}$ converges to x with respect to $\|\cdot\|_\infty$.

Given any $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that for all $k \geq N(\varepsilon)$

$$\max_{1 \leq i \leq n} |x_i^{(k)} - x_i| = \|x^{(k)} - x\|_\infty < \varepsilon$$

$$\Rightarrow |x_i^{(k)} - x_i| < \varepsilon \text{ for each } i$$

$$\Rightarrow \lim_{k \to \infty} x_i^{(k)} = x_i \text{ for each } i.$$

Now Suppose $\lim_{k \to \infty} x_i^{(k)} = x_i$ for every $i = 1, 2, \ldots, n$

Given any $\varepsilon > 0$, let $N_i(\varepsilon)$ for each i represent an integer with the property that

$$|x_i^{(k)} - x_i| < \varepsilon \quad \text{whenever } k \geq N_i(\varepsilon)$$

Let $N(\varepsilon) \equiv \max_i N_i(\varepsilon)$.

If $k \geq N(\varepsilon)$, then $|x_i^{(k)} - x_i| < \varepsilon$ for each i and

$$\max_{1 \leq i \leq n} |x_i^{(k)} - x_i| = \|x^{(k)} - x\|_\infty < \varepsilon$$

$$\Rightarrow \{x^{(k)}\} \text{ converges to } x.$$

Ex. Prove that $x^{(k)} = \left(\frac{1}{k}, 1+e^{1-k}, -2/k^2\right)^t$ (with respect to $\|\cdot\|_\infty$) is convergent, and find the limit of the sequence.

$$\lim_{k\to\infty} \frac{1}{k} = 0, \quad \lim_{k\to\infty} 1+e^{1-k} = 1, \quad \lim_{k\to\infty} -2/k^2 = 0$$

$\therefore$ The theorem implies that the sequence $\{x^{(k)}\}$ converges to $(0, 1, 0)^t$ with respect to $\|\cdot\|_\infty$.

Convergence with respect to the l_2 norm is complicated to check. Instead, we will use the following theorem:

Thm. For each $x \in \mathbb{R}^n$

$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\,\|x\|_\infty$$

Proof. Let x_j be a coordinate of x s.t.

$$\|x\|_\infty = \max_{1\le i\le n} |x_i| = |x_j|$$

Then $\|x\|_\infty^2 = |x_j|^2 = x_j^2 \le \sum_{i=1}^n x_i^2$

$$\le \sum_{i=1}^n x_j^2$$

$$= n\, x_j^2$$

$$= n\, \|x\|_\infty^2$$

Thus, $\|x\|_\infty \le \left\{\sum_{i=1}^n x_i^2\right\}^{1/2} = \|x\|_2 \le \sqrt{n}\, \|x\|_\infty$

∎

Ex. Prove that $x^{(n)} = \left(\frac{1}{n}, 1+e^{1-n}, -2/n^2\right)^t$ is convergent with respect to $\|\cdot\|_2$.

We have $0 \le \|x^{(n)} - x\|_2 \le \sqrt{3}\, \|x^{(n)} - x\|_\infty$

$$\lim_{n\to\infty} \|x^{(n)} - x\|_\infty = 0 \Rightarrow \lim_{n\to\infty} \|x^{(n)} - x\|_2 = 0$$

So $\{x^{(n)}\}$ converges to x wrt $\|\cdot\|_2$.

Indeed, it can be shown that all norms on $\mathbb{R}^n$ are equivalent with respect to Convergence.

ie If $\|\cdot\|_a$ and $\|\cdot\|_b$ are any two norms on $\mathbb{R}^n$ and $\{x^{(k)}\}_{k=1}^{\infty}$ has the limit x wrt $\|\cdot\|_a$ then $\{x^{(k)}\}_{k=1}^{\infty}$ also has the limit x wrt $\|\cdot\|_b$.

We will also use the notion of distance between matrices.

Defn. A matrix norm on the set of all $n \times n$ matrices is a real valued function $\|\cdot\|$ defined on this set satisfying for all $n \times n$ matrices A and B and all real numbers α:

i/ $\|A\| \geq 0$

ii/ $\|A\| = 0 \Leftrightarrow A = 0$

iii/ $\|\alpha A\| = |\alpha| \, \|A\|$

iv/ $\|A + B\| \leq \|A\| + \|B\|$

v/ $\|AB\| \leq \|A\| \, \|B\|$.

Defn A distance between $n \times n$ matrices A and B wrt a matrix norm $\|\cdot\|$ is

$$\|A - B\|.$$

Thm. If $\|\cdot\|$ is a vector norm on $\mathbb{R}^n$, then

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

is a matrix norm.

This is called the natural or induced matrix norm associated with the vector norm.

The following result gives a bound on the value of $\|Ax\|$:

Thm. For any vector $x \neq 0$, matrix A, and any natural norm $\|\cdot\|$ we have

$$\|Ax\| \leq \|A\| \cdot \|x\|.$$

Proof. For $x \neq 0$, $\left\| x/\|x\| \right\| = 1$.

$$\Rightarrow \left\| A\left(\frac{x}{\|x\|}\right) \right\| \leq \|A\|.$$

But $x \neq 0 \Rightarrow A\left(\frac{x}{\|x\|}\right) = \frac{1}{\|x\|} Ax.$

Now $$\frac{1}{\|x\|}\|Ax\| = \left\| \frac{1}{\|x\|} Ax \right\| = \left\| A\left(\frac{x}{\|x\|}\right) \right\| \leq \|A\|$$

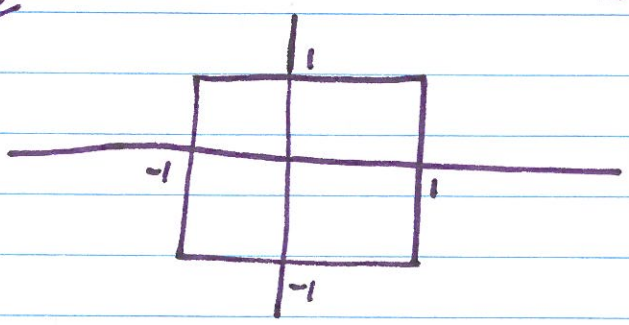
$$\Rightarrow \|Ax\| \leq \|A\| \|x\|$$ ∎

Geometrically, a natural norm describes how the matrix stretches unit vectors relative to that norm.

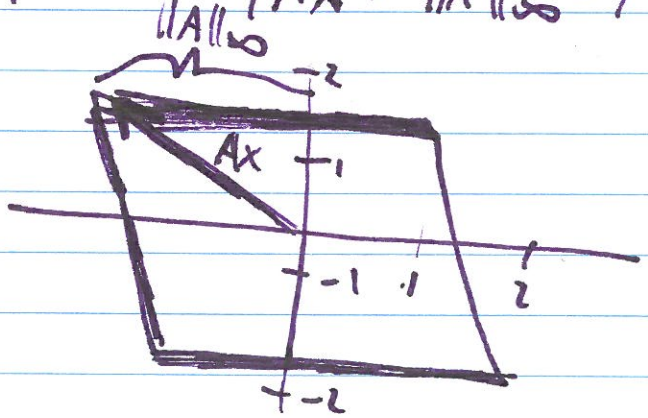
The largest amount of stretch is the norm of the matrix.

ex

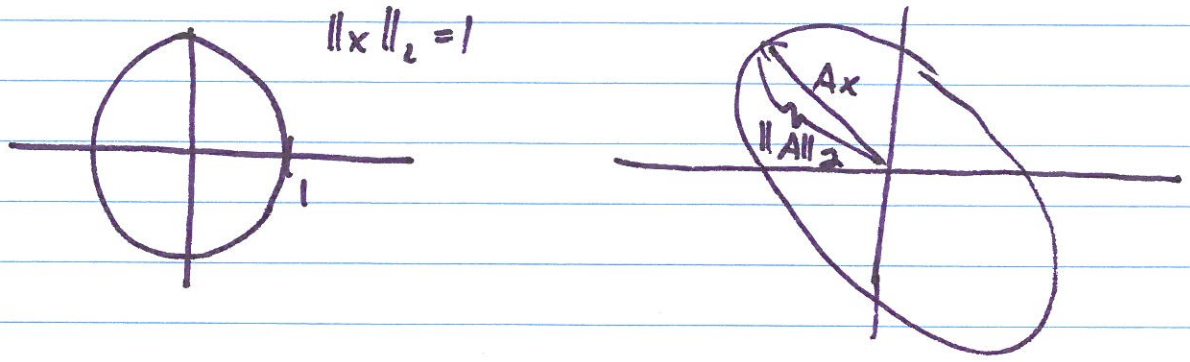
$\{x \in \mathbb{R}^2 : \|x\|_\infty = 1\}$



$\{Ax : \|x\|_\infty = 1\}$



$$\|A\|_\infty = \max_{\|x\|_\infty = 1} \|Ax\|_\infty$$



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

As it turns out, computing the ∞-norm of a matrix is straightforward:

Thm. If $A = (a_{ij})$ is an $n \times n$ matrix then

$$\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$$

Ex Find $\|A\|_\infty$ where $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$.

$$\sum_{j=1}^{n} |a_{1j}| = |2| + |-1| + |0| = 3$$

$$\sum_{j=1}^{n} |a_{2j}| = |-1| + |2| + |-1| = 4$$

$$\sum_{j=1}^{n} |a_{3j}| = |0| + |-1| + |2| = 3$$

$$\Rightarrow \|A\|_\infty = \max\{3, 4, 3\} = 4.$$

Eigenvalues and Eigenvectors

To calculate the l_∞-norm of a matrix we did not need to directly apply to definition.

This is also true for the l_2-norm, however we will need to introduce eigenvalues and eigenvectors to apply this technique.

First we will need the following definition:

Defn If A is a square matrix, the polynomial defined by

$$p(\lambda) = \det(A - \lambda I)$$

is called the characteristic polynomial of A.

It is easily shown that p is an n^{th} degree polynomial.

Now we can introduce eigenvalues and eigenvectors.

Defn. If p is the characteristic polynomial of the matrix A, the zeros of p are called eigenvalues, or characteristic values of the matrix A.

If λ is an eigenvalue of A and $x \neq 0$ has the property that $(A - \lambda I)x = 0$ then x is called an eigenvector, or characteristic vector of A corresponding to the eigenvalue λ.

Ex Compute the eigenvalues and associated eigenvectors of

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Soln. The eigenvalues are determined by solving

$$0 = \det \begin{bmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{bmatrix}$$

$$= (2-\lambda) \det \begin{bmatrix} 2-\lambda & 0 \\ 0 & 3-\lambda \end{bmatrix} - \det \begin{bmatrix} 1 & 0 \\ 0 & 3-\lambda \end{bmatrix}$$

$$\Rightarrow 0 = (2-\lambda)(2-\lambda)(3-\lambda) - (3-\lambda)$$

$$= (3-\lambda)(\lambda^2 - 4\lambda + 3)$$

$$= -(\lambda-3)^2(\lambda-1)$$

$\Rightarrow$ the eigenvalues are

$$\lambda_1 = \lambda_2 = 3 \quad \text{and} \quad \lambda_3 = 1.$$

To determine eigenvectors associated with the eigenvalue $\lambda = 3$ we solve the system

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2-3 & 1 & 0 \\ 1 & 2-3 & 0 \\ 0 & 0 & 3-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_1 + x_2 \\ x_1 - x_2 \\ 0 \end{bmatrix}$$

This implies that $x_1 = x_2$ and that x_3 is arbitrary.

Two linearly independent choices for the eigenvectors associated with the double eigenvalue $\lambda = 3$ are

$$x_1 = (1, 1, 0)^t, \quad x_2 = (1, 1, 1)^t.$$

The eigenvector associated with the eigenvalue $\lambda = 1$ must satisfy

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2-1 & 1 & 0 \\ 1 & 2-1 & 0 \\ 0 & 0 & 3-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 + x_2 \\ 2x_3 \end{bmatrix}$$

This implies that we must have $x_1 = -x_2$ and that $x_3 = 0$.

One choice for the eigenvector associated with the eigenvalue $\lambda = 1$ is

$$x_3 = (1, -1, 0)^t.$$

Notice that if x is an eigenvector associated with the eigenvalue λ, then $Ax = \lambda x$, so the matrix takes the vector into a scalar multiple of itself.

Geometrically, if λ is real, A has the effect of stretching (or shrinking) x by a factor of λ.

Back to finding the l_2-norm of a matrix:

We will need the following definition

Defn: The spectral radius $\rho(A)$ of a matrix A is defined by

$$\rho(A) = \max |\lambda| \quad \text{where } \lambda \text{ is an eigenvalue of } A.$$

Ex. $\rho\left(\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\right) = \max\{|3|, |3|, |1|\}$

$= 3.$

And now we can consider the following

Thm. If A is an $n \times n$ matrix then

(i) $\|A\|_2 = [\rho(A^t A)]^{1/2}$

(ii) $\rho(A) \leq \|A\|$ for any natural norm $\|\cdot\|$.

Proof (ii) Suppose λ is an eigenvalue of A with eigenvector x where $\|x\| = 1$.

$$\Rightarrow Ax = \lambda x.$$

$$|\lambda| = |\lambda| \cdot \|x\| = \|\lambda x\| = \|Ax\|$$

$$\leq \|A\| \, \|x\|$$

any natural norm.

$$= \|A\|.$$

Thus $\rho(A) = \max |\lambda| \leq \|A\|$.

Ex Find $\|A\|_2$ where $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

$$A^tA = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 4 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

To calculate $\rho(A^tA)$ we need the eigenvalues of A^tA

$$0 = \det(A^tA - \lambda I)$$

$$= \det\begin{pmatrix} 5-\lambda & 4 & 0 \\ 4 & 5-\lambda & 0 \\ 0 & 0 & 9-\lambda \end{pmatrix}$$

$$= (5-\lambda)(5-\lambda)(9-\lambda) - (9-\lambda)16$$

$$= (9-\lambda)(25 - 10\lambda + \lambda^2 - 16)$$

$$= (9-\lambda)(9 - 10\lambda + \lambda^2)$$

$$= (9-\lambda)(9-\lambda)(1-\lambda)$$

$\therefore \lambda = 1$ or $\lambda = 9$.

So $\|A\|_2 = \sqrt{\rho(A^tA)} = \sqrt{\max\{1, 9, 9\}} = \sqrt{9} = 3$

When we use iterative matrix techniques, we will want to know when powers of a matrix become small.

Defn We call an $n \times n$ matrix A convergent if

$$\lim_{k \to \infty} (A^k)_{ij} = 0 \quad \text{for each } i, j.$$

Ex Consider $A = \begin{bmatrix} 1/2 & 0 \\ 16 & 1/2 \end{bmatrix}$

$$A^2 = \begin{bmatrix} 1/4 & 0 \\ 16 & 1/4 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} 1/8 & 0 \\ 12 & 1/8 \end{bmatrix}$$

$$A^4 = \begin{bmatrix} 1/16 & 0 \\ 8 & 1/16 \end{bmatrix}$$

$$\vdots$$

$$A^k = \begin{bmatrix} \frac{1}{2^k} & 0 \\ p_k & \frac{1}{2^k} \end{bmatrix}$$

where $p_k \equiv \begin{cases} 16 & \text{if } k = 1 \\ 16/2^{k-1} + \frac{1}{2} p_{k-1} & \text{otherwise.} \end{cases}$

Since $\lim_{n\to\infty} \left(\frac{1}{2n}\right) = 0$

$\lim_{n\to\infty} p_n = 0$ (Why?)

A is a convergent matrix.

Notice that this convergent matrix has a spectral radius less than 1.

This generalizes:

Thm. The following statements are equivalent.

(i) A is a convergent matrix.

(ii) $\rho(A) < 1$

(iii) $\lim_{n\to\infty} A^n x = 0$ for every x.

(iv) $\lim_{n\to\infty} \|A^n\| = 0$ for all natural norms

Iterative Techniques for Solving Linear Systems

In problems where the matrix A contains a high proportion of zeros, iterative techniques are often used to solve the system

$$Ax = b$$

since they preserve the zero structure of the matrix.

Iterative techniques convert the system

$$Ax = b$$

into an equivalent system of the form

$$x = Tx + c$$

T: fixed matrix; c: vector c.

An initial vector $x^{(0)}$ is selected, and then a sequence of approximate solution vectors is generated:

$$x^{(k)} = T x^{(k-1)} + c$$

Iterative techniques are rarely used in very small systems (ie when n^3 is small) in these cases iterative methods may be slower since they require several iterations to obtain the desired accuracy.

IDEA: It is possible to "split" the matrix A:

$$Ax = b$$

$$[M + (A - M)]x = b$$

$$Mx = b + (M - A)x$$

$$x = (I - M^{-1}A)x + M^{-1}b.$$

Iteration becomes

$$x^{(k+1)} = \underbrace{(I - M^{-1}A)}_{\text{amplification matrix} \equiv T} x^{(k)} + \underbrace{M^{-1}b}_{\equiv c}$$

$$x^{(k+1)} = T x^{(k)} + c.$$

How to choose M?

Want i/ M easy to "invert"

ii/ M "close to A" in the sense that $\rho(T)$ is small.

Ex Let $M = D \equiv \begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$

This gives the

Jacobi Iterative Method.

In the text's notation

$$A = D - L - U$$

$$\begin{bmatrix} a_{11} & & 0 \\ & a_{22} & \\ 0 & & \ddots \\ & & & a_{nn} \end{bmatrix} \quad \begin{bmatrix} 0 & & & \\ -a_{21} & 0 & & 0 \\ -a_{31} & -a_{32} & \ddots & \\ \vdots & & & 0 \\ -a_{n1} & & -a_{n,n-1} & \end{bmatrix} \quad \begin{bmatrix} 0 & -a_{12} & -a_{13} & \cdots & -a_{1n} \\ & 0 & -a_{23} & & \\ & & \ddots & & \vdots \\ & & & & -a_{n-1,n} \\ & & & 0 & \end{bmatrix}$$

$$Ax = b$$

$$(D - L - U)x = b$$

$$Dx = (L+U)x + b$$

$$x = D^{-1}(L+U)x + D^{-1}b.$$

Which results in the iteration

$$x^{(k+1)} = \underbrace{D^{-1}(L+U)}_{\equiv T} x^{(k)} + \underbrace{D^{-1}b}_{\equiv c}$$

Eg solve $\begin{bmatrix} 10 & -1 & 2 & 0 \\ -1 & 11 & -1 & 3 \\ 2 & -1 & 10 & -1 \\ 0 & 3 & -1 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 25 \\ -11 \\ 15 \end{bmatrix}$

by Jacobi's Method.

$$D = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 11 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}, \quad D^{-1} = \begin{bmatrix} \frac{1}{10} & 0 & 0 & 0 \\ 0 & \frac{1}{11} & 0 & 0 \\ 0 & 0 & \frac{1}{10} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{bmatrix}$$

$$L+U = \begin{bmatrix} 0 & 1 & -2 & 0 \\ 1 & 0 & 1 & -3 \\ -2 & 1 & 0 & 1 \\ 0 & -3 & 1 & 0 \end{bmatrix}$$

$$\underbrace{T}_{D^{-1}(L+U)} = \begin{bmatrix} 0 & \frac{1}{10} & -\frac{2}{10} & 0 \\ \frac{1}{11} & 0 & \frac{1}{11} & -\frac{3}{11} \\ -\frac{2}{10} & \frac{1}{10} & 0 & \frac{1}{10} \\ 0 & -\frac{3}{8} & \frac{1}{8} & 0 \end{bmatrix}$$

$$\underbrace{C}_{D^{-1}\begin{bmatrix} 6 \\ 25 \\ -11 \\ 15 \end{bmatrix}} = \begin{bmatrix} 6/10 \\ 25/11 \\ -11/10 \\ 15/8 \end{bmatrix}$$

36.76

Take $x^{(0)} = (0, 0, 0, 0)^t$.

$$x^{(1)} = Tx^{(0)} + c = c = \begin{pmatrix} 6/10 \\ 25/11 \\ -11/10 \\ 15/8 \end{pmatrix}$$

$$x^{(2)} = Tx^{(1)} + c = \begin{pmatrix} 1.0473 \\ 1.7159 \\ -0.8052 \\ 0.8852 \end{pmatrix}$$

$\vdots$

$$x^{(9)} = \begin{pmatrix} 0.9997 \\ 2.0004 \\ -1.0004 \\ 1.0006 \end{pmatrix}$$

$$x^{(10)} = \begin{pmatrix} 1.0001 \\ 1.9998 \\ -0.9998 \\ 0.9998 \end{pmatrix}$$

Comments on Jacobi's Method

$$x^{(n+1)} = D^{-1}(L+U)x^{(n)} + D^{-1}b$$

① The algorithm requires that $a_{ii} \neq 0$ for each i. If one of the $a_{ii} = 0$, and the system is nonsingular, then a reordering of the equations can be performed so that no $a_{ii} = 0$.

② To speed convergence the equations should be arranged so that $|a_{ii}|$ is as large as possible.

③ A possible stopping criterion is to iterate until

$$\frac{\|x^{(n)} - x^{(n-1)}\|}{\|x^{(n)}\|} < \varepsilon$$

If we write out Jacobi's Method

$$x^{(k+1)} = D^{-1}(L+U)\,x^{(k)} + D^{-1}b$$

we find that

$$x_i^{(k+1)} = \frac{\sum_{j \neq i} \left(-a_{ij}\, x_j^{(k)}\right) + b_i}{a_{ii}}$$

Notice that to compute $x_i^{(k+1)}$ the components $x_i^{(k)}$ are used.

But for $i > 1$, $x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_{i-1}^{(k+1)}$

have already been computed and are likely better approximations to the actual solutions than

$$x_1^{(k)}, x_2^{(k)}, \ldots, x_{i-1}^{(k)}.$$

So it seems reasonable to compute with these most recently computed values...

ie $$x_i^{(k+1)} = \frac{-\sum_{j=1}^{i-1}\left(a_{ij}x_j^{(k+1)}\right) - \sum_{j=i+1}^{n}\left(a_{ij}x_j^{(k)}\right) + b_i}{a_{ii}}$$

This is called the Gauss-Seidel iterative technique and it also has a matrix formulation with $M \equiv (D-L)$:

$$Ax = b$$

$$(D-L-U)x = b.$$

$$(D-L)x = Ux + b$$

$$x = (D-L)^{-1}Ux + (D-L)^{-1}b.$$

$\Rightarrow$ iteration becomes

$$x^{(k+1)} = \underbrace{(D-L)^{-1}U}_{T_g}x^{(k)} + \underbrace{(D-L)^{-1}b}_{C_g}.$$

notice that $D-L$ is lower triangular. It is invertible $\Leftrightarrow$ each $a_{ii} \neq 0$.

We want to study the convergence of general iteration techniques

$$x^{(k)} = T x^{(k-1)} + C$$

We will need the following Lemma:

Lemma: If the spectral radius $\rho(T)$ satisfies $\rho(T) < 1$ then $(I - T)^{-1}$ exists and

$$(I - T)^{-1} = I + T + T^2 + \cdots$$

And we will prove the following theorem:

Thm: For any $x^{(0)} \in \mathbb{R}^n$, the sequence $\{x^{(k)}\}_{k=0}^{\infty}$ defined by

$$x^{(k)} = T x^{(k-1)} + C$$

converges to the unique solution of

$$x = Tx + C$$ if and only if $\rho(T) < 1$.

Proof $\Leftarrow$ assume $\rho(T) < 1$.

$$x^{(k)} = Tx^{(k-1)} + c$$

$$= T(Tx^{(k-2)} + c) + c$$

$$= T^2 x^{(k-2)} + (T+I)c$$

$$= \vdots$$

$$= T^k x^{(0)} + (T^{k-1} + \cdots + T + I)c$$

Since $\rho(T) < 1$, the matrix T is convergent and

$$\lim_{k \to \infty} T^k x^{(0)} = 0.$$

The Lemma implies that

$$\lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} T^k x^{(0)} + \lim_{k\to\infty} \left(\sum_{j=0}^{k-1} T^j \right) c$$

$$= 0 + (I - T)^{-1} c$$

$\Rightarrow \{x^{(k)}\}$ converges to the unique solution of $x = Tx + c$.

ie $(I - T)x = c \Rightarrow x = (I-T)^{-1}c$.

forward result.

Let $x^{(0)} \in \mathbb{R}^n$ be arbitrary.

Given $\{x^{(k)}\}_{k=0}^{\infty}$ converges to $x = (I-T)^{-1}c$

where $x^{(k)} = Tx^{(k-1)} + c$.

From our previous result,

$$x^{(k)} = T^k x^{(0)} + (I-T)^{-1}c.$$

$$x^{(k)} = T^k x^{(0)} + x.$$

$$\lim_{k\to\infty} x - x^{(k)} = \lim_{k\to\infty} -T^k x^{(0)}$$

$$T^k x^{(0)} = 0.$$

but $x^{(0)}$ was arbitrary so T is a convergent matrix & $\rho(T) < 1$.

This allows us to derive some related results on the rates of convergence.

Corollary: If $\|T\| < 1$ for any natural matrix norm and c is a given vector, then the sequence $\{x^{(k)}\}_{k=0}^{\infty}$ defined by

$$x^{(k)} = Tx^{(k-1)} + c$$

converges for any $x^{(0)} \in \mathbb{R}^n$ to a vector $x \in \mathbb{R}^n$ and the following error bounds hold:

(i) $\|x - x^{(k)}\| \leq \|T\|^k \|x^{(0)} - x\|$

(ii) $\|x - x^{(k)}\| \leq \dfrac{\|T\|^k}{1 - \|T\|} \|x^{(1)} - x^{(0)}\|$

Note however that $\rho(A) \leq \|A\|$ for any natural norm. In practice

$$\|x - x^{(k)}\| \approx \rho(T)^k \|x^{(0)} - x\|$$

so it is desirable to have $\rho(T)$ as small as possible.

Some results for Jacobi and Gauss-Seidel Methods.

Thm. If A is strictly diagonally dominant, then for any choice of $x^{(0)}$, both the Jacobi and Gauss-Seidel methods give sequences $\{x^{(k)}\}_{k=0}^{\infty}$ that converge to the unique solution of $Ax = b$.

No general results exist to tell which of the two methods will converge more quickly, but the following result applies in a variety of examples:

Thm. Stein Rosenberg.

If $a_{ij} \leq 0$ for each $i \neq j$ and $a_{ii} > 0$ for each $i = 1, 2, \ldots, n$ then exactly one of the following holds.

a. $0 \leq \rho(T_g) < \rho(T_j) < 1.$
b. $1 < \rho(T_j) < \rho(T_g)$
c. $\rho(T_j) = \rho(T_g) = 0$
d. $\rho(T_j) = \rho(T_g) = 1.$

Notice that if one method converges, both do & Gauss' Seidel converges faster.

Otherwise if one method diverges, both do. The divergence for Gauss Seidel is more pronounced.

This result only holds when $a_{ij} \leq 0$ / $i \neq j$
$a_{ii} > 0$

Another simple, but useful iterative method is

Successive Over Relaxation (SOR).

To define, suppose $\tilde{x}^{(k+1)}$ is the iterate from Gauss-Seidel using $x^{(k)}$

The $(k+1)^{st}$ iterate of SOR is defined by

$$x^{(k+1)} = \omega \tilde{x}^{(k+1)} + (1-\omega) x^{(k)}$$

where $1 < \omega < 2$. It can be difficult to select ω optimally. Indeed, the answer to this question is not known for general $n \times n$ linear systems.

However, we do have the following results:

Thm. (Kahan) If $a_{ii} \neq 0$ for each i then

$$\rho(T_{SOR}) \geq |\omega - 1|$$

$\Rightarrow$ SOR can converge only if $0 < \omega < 2$.

Thm. (Ostrowski-Reich) If A is a positive definite matrix and $0 < \omega < 2$ then the SOR method converges for any choice of initial approximate vector $x^{(0)}$.

Thm. If A is positive definite and tridiagonal, then $\rho(T_g) = [\rho(T_j)]^2 < 1$ and the optimal choice of ω for the SOR method is

$$\omega = \frac{2}{1 + \sqrt{1 - \rho(T_j)^2}}$$

With this choice of ω, we have $\rho(T_{SOR}) = \omega - 1$.