Iterative methods :

Notation and a brief background

- **Mathematical background: matrices, inner products and norms**
- **linear systems of equations**
- **Iterative processes**

Notation & Review of some linear algebra concepts

➤ The set of all linear combinations of a set of vectors $G = \{a_1, a_2, \ldots, a_q\}$ of $\mathbb{R}^n$ is a vector subspace called the linear span of G. Notation

$$\text{span}(G), \text{ span } \{a_1, a_2, \ldots, a_q\}$$

➤ If the a_i's are linearly independent, then each vector of $\mathbf{span}\{G\}$ admits a unique expression as a linear combination of the a_i's. The set G is then called a *basis*

➤ Recall: A matrix represents a linear mapping between two vector spaces of finite dimension n and m.

Transposition: If $A \in \mathbb{R}^{m\times n}$ then its transpose is a matrix $C \in \mathbb{R}^{n\times m}$ with entries

$$c_{ij} = a_{ji}, i = 1, \ldots, n,\ j = 1, \ldots, m$$

Notation : A^T.

Transpose Conjugate: for complex matrices, the **transpose conjugate** matrix denoted by A^H is more relevant: $A^H = \bar{A}^T = \overline{A^T}$.

- We consider now only square matrices ($m = n$).
- Spectral radius = The maximum modulus of the eigenvalues

$$\boxed{\rho(A) = \max_{\lambda\in\lambda(A)} |\lambda|.}$$

- Recall: $\lim_{k\to\infty} A^k = 0$ iff $\rho(A) < 1$.

➤ Trace of A = sum of diagonal elements of A.

$$\mathbf{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

➤ $\mathbf{tr}(A)$ = sum of all the eigenvalues of A counted with their multiplicities.

➤ Recall that $\mathbf{det}(A)$ = product of all the eigenvalues of A counted with their multiplicities.

$\boxed{\textit{Example:}}$: Trace, spectral radius, and determinant of the matrix:

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

Range and null space

- Range: $\mathbf{Ran}(A) = \{Ax \mid x \in \mathbb{R}^n\}$
- Null Space: $\mathbf{Null(A)} = \{x \in \mathbb{R}^n \mid Ax = \mathbf{0}\ \}$
- Range $=$ linear span of the columns of A
- Rank of a matrix $\mathbf{rank}(A) = \mathbf{dim}(\mathbf{Ran}(A))$
- rank (A) $=$ the number of linearly independent columns of (A) $=$ the number of linearly independent rows of A.
- A is of *full rank* if $\mathbf{rank}(A) = \mathbf{min}\{m, n\}$. Otherwise it is *rank-deficient.*

Rank+Nullity theorem for an $m \times n$ matrix:

$$dim(Ran(A)) + dim(Null(A)) = n$$

Types of matrices (square)

- *Symmetric matrices:* $A^T = A$.
- *Hermitian matrices:* $A^H = A$.
- *Skew-symmetric matrices:* $A^T = -A$.
- *Skew-Hermitian matrices:* $A^H = -A$.
- *Normal matrices:* $A^H A = AA^H$.
- *Nonnegative matrices:* $a_{ij} \geq 0, \; i, j = 1, \ldots, n$ (similar definition for nonpositive, positive, and negative matrices).
- *Unitary matrices:* $Q^H Q = I$.

Note: if Q is unitary then $Q^{-1} = Q^H$.

Inner products and Norms

- Inner product of 2 vectors x and y in $\mathbb{R}^n$:

$$\boxed{x_1y_1 + x_2y_2 + \cdots + x_ny_n \text{ in } \mathbb{R}^n}$$

Notation: (x, y) or y^Tx

- For complex vectors

$$\boxed{(x, y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n \text{ in } \mathbb{C}^n}$$

Note: $(x, y) = y^Hx$

An important property: Given $A \in \mathbb{C}^{m\times n}$ then

$$(Ax, y) = (x, A^Hy) \quad \forall\, x \in \mathbb{C}^n, \forall y \in \mathbb{C}^m$$

Vector norms

Norms are needed to measure lengths of vectors and closeness of two vectors. Examples of use: Estimate convergence rate of an iterative method; Estimate the error of an approximation to a given solution; ...

➤ A vector norm on a vector space $\mathbb{X}$ is a real-valued function on $\mathbb{X}$, which satisfies the following three conditions:

1. $\|x\| \geq 0, \quad \forall\, x \in \mathbb{X}$, and $\|x\| = 0$ iff $x = 0$.
2. $\|\alpha x\| = |\alpha| \|x\|, \quad \forall\, x \in \mathbb{X}, \quad \forall \alpha \in \mathbb{C}$.
3. $\|x + y\| \leq \|x\| + \|y\|, \quad \forall\, x, y \in \mathbb{X}$.

➤ 3. is called the triangle inequality.

Example: Euclidean norm on $\mathbb{X} = \mathbb{C}^n$,

$$\|x\|_2 = (x, x)^{1/2} = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2}$$

➤ Most common vector norms in numerical linear algebra:

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|,$$
$$\|x\|_2 = \left[|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2\right]^{1/2},$$
$$\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|.$$

➤ The Cauchy-Schwartz inequality (important) is:

$$|(x, y)| \leq \|x\|_2 \|y\|_2.$$

Convergence of vector sequences

A sequence of vectors $x^{(k)}$, $k = 1, \ldots, \infty$ converges to a vector x with respect to the norm $\|.\|$ if, by definition,

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

- Important point: because all norms in $\mathbb{R}^n$ are equivalent, the convergence of $x^{(k)}$ w.r.t. a given norm implies convergence w.r.t. any other norm.
- Notation: $\boxed{\lim_{k\to\infty} x^{(k)} = x}$
- Note: $x^{(k)}$ converges to x iff each component $x_i^{(k)}$ of $x^{(k)}$ converges to the corresoponding component x_i of x

Matrix norms

➤ Can define matrix norms by considering $m \times n$ matrices as vectors in $\mathbb{R}^{mn}$. These norms satisfy the usual properties of vector norms, i.e.,

1. $\|A\| \geq 0, \ \forall \ A \ \in \mathbb{C}^{m\times n}$, and $\|A\| = 0$ iff $A = 0$
2. $\|\alpha A\| = |\alpha| \|A\|, \forall \ A \ \in \mathbb{C}^{m\times n}, \ \forall \ \alpha \in \mathbb{C}$
3. $\|A + B\| \leq \|A\| + \|B\|, \ \forall \ A, B \ \in \mathbb{C}^{m\times n}$.

➤ However, these will lack (in general) the right properties for composition of operators (product of matrices).

➤ The case of $\|.\|_2$ yields the Frobenius norm of matrices.

- Given a matrix A in $\mathbb{C}^{m\times n}$, define the set of matrix norms

$$\|A\|_p = \max_{x\in\mathbb{C}^n,\ x\neq 0} \frac{\|Ax\|_p}{\|x\|_p}.$$

- These norms satisfy the usual properties of vector norms (see previous page).
- The matrix norm $\|.\|_p$ is induced by the vector norm $\|.\|_p$.
- Again, important cases are for $p = 1, 2, \infty$.

Consistency

- A fundamental property is consistency

$$\|AB\|_p \le \|A\|_p \|B\|_p.$$

- Consequence: $\boxed{\|A^k\|_p \le \|A\|_p^k}$

- A^k converges to zero if *any* of its p-norms is < 1

- The **Frobenius norm** is defined by

$$\|A\|_F = \left(\sum_{j=1}^{n} \sum_{i=1}^{m} |a_{ij}|^2 \right)^{1/2}.$$

- Same as the 2-norm of the column vector in $\mathbb{C}^{mn}$ consisting of all the columns (respectively rows) of A.

- This norm is also consistent [but not induced from a vector norm]

Important equalities:

$$\|A\|_1 = \max_{j=1,\ldots,n} \sum_{i=1}^{m} |a_{ij}|,$$
$$\|A\|_\infty = \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{ij}|,$$
$$\|A\|_2 = \left[\rho(A^H A)\right]^{1/2} = \left[\rho(AA^H)\right]^{1/2},$$
$$\|A\|_F = \left[Tr(A^H A)\right]^{1/2} = \left[Tr(AA^H)\right]^{1/2}.$$

Positive-Definite Matrices

- A real matrix is said to be positive definite if

$$(Au, u) > 0 \text{ for all } u \neq 0 \; u \in \mathbb{R}^n$$

- Let A be a real positive definite matrix. Then there is a scalar $\alpha > 0$ such that

$$(Au, u) \geq \alpha \|u\|_2^2,$$

- Consider now the case of Symmetric Positive Definite (SPD) matrices.

- Consequence 1: A is nonsingular

- Consequence 2: the eigenvalues of A are (real) positive

A few properties of Symmetric Positive Definite matrices

- Diagonal entries of A are positive. More generally, ...
- Diagonal block $A(k:l, k:l)$, $(k<l)$, is SPD
- For any $n \times k$ matrix X of rank k, the matrix X^TAX is SPD.
- The mapping :

$$x, y \quad \rightarrow \quad (x, y)_A \equiv (Ax, y)$$

is a proper inner product on $\mathbb{R}^n$. The associated norm, denoted by $\|.\|_A$, is called the energy norm:

$$\|x\|_A = (Ax, x)^{1/2} = \sqrt{x^TAx}$$

- A admits the Cholesky factorization $A = LL^T$ where L is lower triangular

Iterative processes for linear systems

In contrast with "direct" methods (Gaussian Elimination) iterative methods compute a sequence of approximations $\boldsymbol{x}^{(k)}$ to the solution $\boldsymbol{x}$. Ideally, a good approximation is obtained in a few iterations of the process. Convergence measured by some norm.

Questions which arise:

- Why not use a direct method [always works!]
- Under which condition (s) will the method converge?
- When to stop?
- Can we estimate costs?

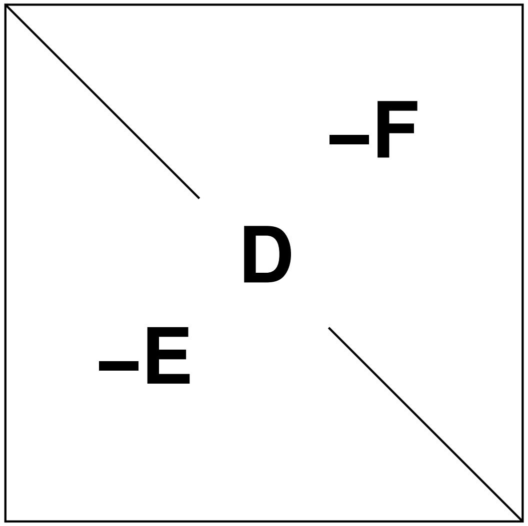
Basic relaxation techniques

- Relaxation methods: Jacobi, Gauss-Seidel, SOR
- Basic convergence results
- Optimal relaxation parameter for SOR
- See Chapter 4 of text for details.

Basic relaxation schemes

➤ *Relaxation schemes:* methods that modify one component of current approximation at a time

➤ Based on the decomposition $A = D - E - F$ with: $D = \text{diag(A)}$, $-E =$ strict lower part of A and $-F =$ its strict upper part.



Gauss-Seidel iteration for solving $Ax = b$:

➤ corrects j-th component of current approximate solution, to zero the $j - th$ component of residual for $j = 1, 2, \cdots, n$.

- Gauss-Seidel iteration can be expressed as:

$$\boxed{(D - E)x^{(k+1)} = Fx^{(k)} + b}$$

Can also define a backward Gauss-Seidel Iteration:

$$(D - F)x^{(k+1)} = Ex^{(k)} + b$$

and a Symmetric Gauss-Seidel Iteration: forward sweep followed by backward sweep.

Over-relaxation is based on the decomposition:

$$\omega A = (D - \omega E) - (\omega F + (1 - \omega)D)$$

$\longrightarrow$ successive overrelaxation, (SOR):

$$\boxed{(D - \omega E)x^{(k+1)} = [\omega F + (1 - \omega)D]x^{(k)} + \omega b}$$

Iteration matrices

Jacobi, Gauss-Seidel, SOR, & SSOR iterations are of the form

$$x^{(k+1)} = Mx^{(k)} + f$$

$$\begin{aligned}
M_{Jac} &= D^{-1}(E+F) = I - D^{-1}A \\
M_{GS} &= (D-E)^{-1}F = I - (D-E)^{-1}A \\
M_{SOR} &= (D-\omega E)^{-1}(\omega F + (1-\omega)D) \\
&= I - (\omega^{-1}D - E)^{-1}A \\
M_{SSOR} &= I - \omega(2-\omega)(D-\omega F)^{-1}D(D-\omega E)^{-1}A
\end{aligned}$$

General convergence result

Consider the iteration: $x^{(k+1)} = Gx^{(k)} + f$

(1) Assume that $\rho(G) < 1$. Then $I - G$ is non-singular and G has a fixed point. Iteration converges to a fixed point for any f and $x^{(0)}$.

(2) If iteration converges for any f and $x^{(0)}$ then $\rho(G) < 1$.

Example: Richardson's iteration

$$x^{(k+1)} = x^{(k)} + \alpha(b - Ax^{(k)})$$

✍ Assume $\Lambda(A) \subset \mathbb{R}$. When does the iteration converge?

A few well-known results

➤ Jacobi and Gauss-Seidel converge for diagonal dominant matrices, i.e., matrices such that

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, i = 1, \cdots, n$$

➤ SOR converges for $0 < \omega < 2$ for SPD matrices

➤ The optimal ω is known in theory for an important class of matrices called 2-cyclic matrices or matrices with property A.

➤ A matrix has property A if it can be (symmetrically) permuted into a 2×2 block matrix whose diagonal blocks are diagonal.

$$PAP^T = \begin{bmatrix} D_1 & E \\ E^T & D_2 \end{bmatrix}$$

➤ Let A be a matrix which has property A. Then the eigenvalues λ of the SOR iteration matrix and the eigenvalues μ of the Jacobi iteration matrix are related by

$$(\lambda + \omega - 1)^2 = \lambda \omega^2 \mu^2$$

➤ The optimal ω for matrices with property A is given by

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho(B)^2}}$$

where B is the Jacobi iteration matrix.

An observation — Introduction to Preconditioning

➤ The iteration $x^{(k+1)} = Mx^{(k)} + f$ is attempting to solve $(I - M)x = f$. Since M is of the form $M = I - P^{-1}A$ this system can be rewritten as

$$P^{-1}Ax = P^{-1}b$$

where for SSOR, we have

$$P_{SSOR} = (D - \omega E)D^{-1}(D - \omega F)$$

referred to as the SSOR 'preconditioning' matrix.

In other words:

Relaxation iter. $\Longleftrightarrow$ *Preconditioned Fixed Point Iter.*