Solving Systems of Equations

IE 496 Lecture 22

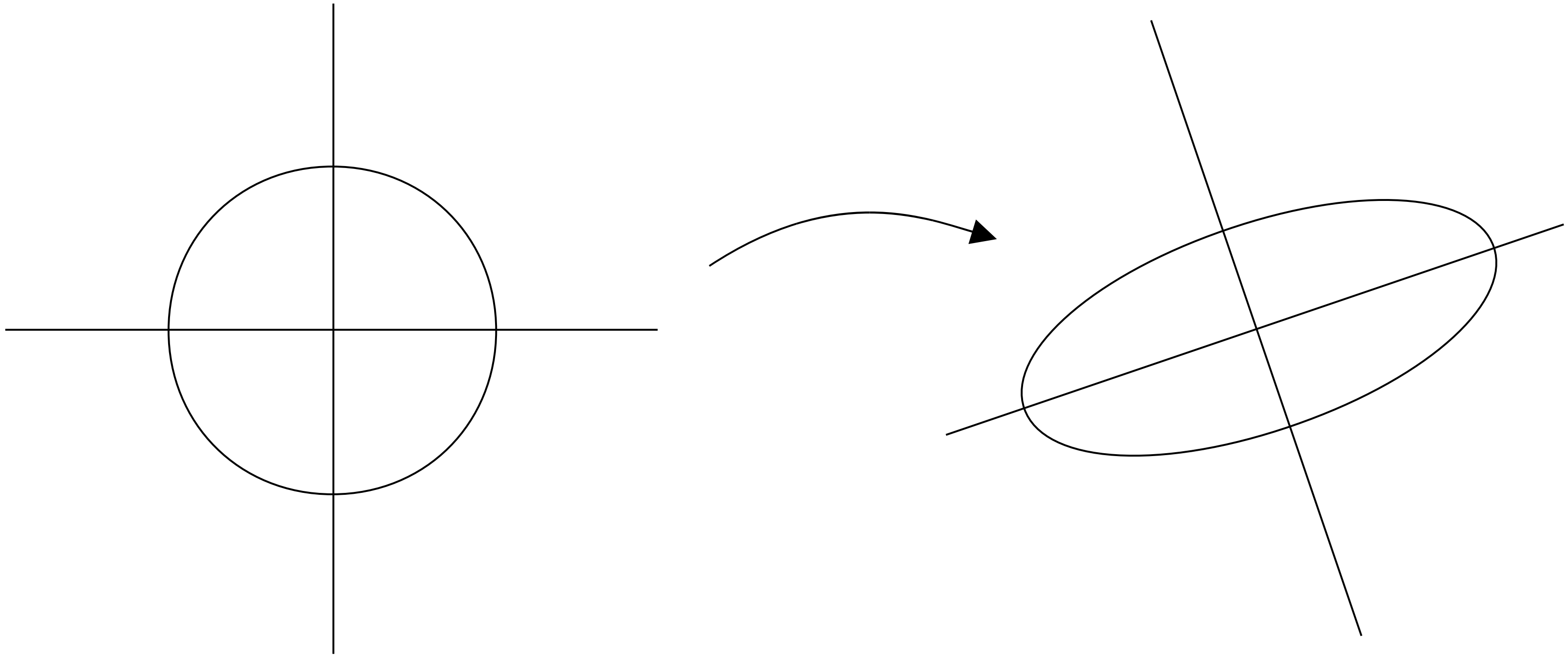
Reading for This Lecture

- Primary
  - Miller and Boxer, Pages 128-134
  - Forsythe and Mohler, Sections 9 and 10

Solving Systems of Equations

- Problem: Given a matrix $A \in \mathbf{R}^{n \times n}$ and a vector $b \in \boldsymbol{R}^n$, we wish to find $x \in \mathbf{R}^n$ such that $Ax = b$.
- Diagonal form of a matrix
  - An othogonal matrix U has the property the $U^T U = UU^T = I$.
  - Given $A \in \mathbf{R}^{n \times n}$, there exist orthogonal matrices U, V such that
    - $U^T A V = D$ where D is a diagonal matrix where
    - diagonal elements of D are $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_r > \mu_{r+1} = \cdots = \mu_n = 0$, and
    - r is the rank of A.
    - μ_i is the non-negative square root of the i^{th} eigenvalue.
  - This is called the *singular value decomposition*.

Importance of the SVD



Effect of multiplying by a matrix

Implications

- Multiplying by A represents a *rotation* and a *scaling* of axes to get from one space to the other.
- μ_i is the non-negative square root of the i^{th} eigenvalue.
- Notice that $||A|| = ||D|| = \mu_1$.
- So the norm of A is the maximum amount any axis gets magnified by A.
- If $r = n$, then we can easily derive the inverse of A.
- Also, $||A^{-1}|| = ||A||^{-1} = 1/\mu_n$.

Condition of a Linear System

- Consider the problem of solving $Ax = b$.
- If we perturb b, how much does the x change?
- $x + \delta x = A^{-1}(b + \delta b) \Rightarrow \delta x = A^{-1}\delta b$
- $\|\delta x\| \leq \|A^{-1}\| \cdot \|\delta b\|$
- $\|\delta x\| \cdot \|b\| \leq \|A\| \cdot \|A^{-1}\| \cdot \|x\| \cdot \|\delta b\|$
- $\|\delta x\|/\|x\| \leq \|A\| \cdot \|A^{-1}\| \cdot \left(\|\delta b\|/\|b\|\right)$
- The *condition number* of a matrix is the quantity
  $\mathrm{cond}(A) = \|A\| \cdot \|A^{-1}\|$

Condition Number

- Note that cond(A) = μ_1/μ_n.
- Hence it is a relative measure of how much distortion A causes to its input.
- It is also a measure of how much the inaccuracies in b get multiplied in x when solving systems $Ax = b$.
- If b is the result of a previous calculation, then $\|\delta b\|/\|b\|$ is *at best* equal to u (machine epsilon).
- The inaccuracies in x will then be *at best* $u \cdot \text{cond}(A)$.

Interpretation

- Orthogonal matrices have a norm of 1 and hence don't cause any scaling or distortion.
- Singular matrices have at least one singular value equal to 0 and hence have a norm of "infinity".
- "Nearly singular" matrices are the ones that cause problems.
- These are ones that have singular values "relatively close" to zero.

Gaussian Elimination

- Standard row operations
  - Interchange rows
  - Multiply rows by a scalar
  - Subtract a multiple of row j from row i
- Standard algorithm
  - Elimination Phase
  - Back-substitution Phase

Gaussian Elimination

- Elimination Phase
  - For $i = 1$ to n
    - Exchange row i with row $j > i$ to ensure $A_{ii} \neq 0$ (if not possible, STOP).
    - Scale row i so that $A_{ii} = 1$
    - For $j = i+1$ to n
      - Subtract A_{ij} times row i from row j so that $A_{ij} = 0$
- Back Substitution Phase
  - For $i = n$ to 1
    - For $j = i-1$ to 1
      - Subtract A_{ij} times row i from row j so that $A_{ij} = 0$

The LU Factorization

- The *LU decomposition*
  - Assume $\det(A_k) \neq 0 \; \forall k$
  - $\exists$ a lower triangular matrix L with 1's on the diagonal, and
  - an upper triangular matrix U such that
  - $A = LU$
- With an *LU factorization*, can solve the system $Ax = b$
- Solve $Ly = b$ (elimination phase)
- Solve $Ux = y$ (back substitution phase)
- Hence, we see the relationship to Gaussian Elimination.

Calculating an LU Factorization

- The LU factorization can be computed "in-place" (sort of).
- Row interchanges can be represented by *permutation matrices*.
- Elimination operations can be represented by *eta matrices*.
- The eta matrices can be stored compactly as elimination proceeds.
- In the end, you have an *LU* decomposition.

Solving with Multiple RHS's

- Suppose we wish to solve the system $Ax = b$ with multiple RHS vectors.
- Calculate an LU factorization.
- Use it to solve the system with various RHS's.
- Avoid computing A^{-1}
  - Takes more computation (takes longer)
  - More round-off error
  - Usually completely dense

More On Row Interchanges

- Bad Example
- Partial Pivoting Strategy
  - Take the pivot element to be the largest element (in absolute value) in the column
- Complete Pivoting Strategy
  - Take the pivot element to be the largest element (in absolute value) in the whole matrix
- Using these strategies, we can limit round-off error
- Roughly, we will obtain x such that $(A + \delta A)x = b$ and the entries of δA are $O(nu)$.

Parallel Gaussian Elimination

- PRAM with n^2 processors

- Mesh with n^2 processors