Lecture 02 - Solving Linear Systems

May 7, 2025

Outline

1. High-Level Motivation
2. Solving Linear Systems
   1. LU factorizations
      1. Complexity
   2. Symmetric Systems
   3. Positive Definite Systems

High-Level Motivation

Applications We Will Study

1. Solving Linear Systems
   1. heat conduction (including difference equations)
   2. image de-noising
2. Least Squares
   1. motivated by statistical methods
3. Eigenvalues / Eigenvectors
   1. image segmentation
4. Singular Value Decompositions
   1. image compression

Solving Linear Systems

How To Compute $x = A^{-1}b$: In numerical linear algebra, we never compute A^{-1} in order to compute $A^{-1}b$. Instead we compute x as the solution of $Ax = b$, via Gaussian elimination.

Big Picture of Gaussian Elimination

$$
\underbrace{\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{bmatrix}}_{A} \to \underbrace{\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix}}_{A^{(1)}} \to
$$

$$
\underbrace{\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & x & x \end{bmatrix}}_{A^{(2)}} \to \underbrace{\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{bmatrix}}_{A^{(3)}}
$$

Solving Linear Systems

GE Algorithm

for i = 1, 2, ..., n-1
    for k = i+1, ..., n
        mult = a_{ki} / a_{ii}
        $a_{ki} = 0$ not needed, but helpful for intuition
        for j = i+1, ..., n
            $a_{kj} = a_{kj} -$ mult $*a_{ij}$ update row k
        end
        $b_k = b_k -$ mult $*b_i$ update RHS
    end
end

At the end, $A^{(n-1)}x = b^{(n-1)}$, is solved by back substitution.

Solving Linear Systems - LU Factorizations

Theorem 1

If A can be reduced to RREF without interchanging rows, then there is a unique factorization $A = LU$, where L is lower triangular with 1s on its diagonal (i.e. unit diagonal), and U is upper triangular. Moreover,

$$U = A^{(n-1)}, L = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ mult & & 1 \end{bmatrix}$$

Proof.

See the proof, starting on p144 of *Matrix Analysis and Applied Linear Algebra*, by Carl D. Meyer. □

Important Remark: Not every non-singular $n \times n$ matrix A has an LU-decomposition. E.g. $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$.

Solving Linear Systems - LU Factorizations

Then solving $Ax = b$ is equivalent to solving $LUx = b$. Let $y = Ux$. Then $Ly = b$. So we

1. Solve $Ly = b$ by forward solving, then
2. Solve $Ux = y$ by back solving.

Solving Linear Systems - LU Factorizations

Forward Solve Algorithm

for $i = 1, 2, \ldots, n$

$\quad y_i = b_i$

$\quad$ for $j = 1, 2, \ldots, i-1$

$\qquad y_i = y_i - l_{ij} * y_j \qquad \left(y_i = b_i - \sum_{j=1}^{i-1} l_{ij} y_j \right)$

$\quad$ end

end

Solving Linear Systems - LU Factorizations

Backward Solve Algorithm

for $i = n, \ldots, 1$
    $x_i = y_i$
    for $j = i + 1, \ldots, n$
        $x_i = x_i - u_{ij} * x_j$ $\left(x_i = y_i - \sum_{j=i+1}^{n} u_{ij} x_j\right)$
    end
    $x_i = x_i / u_{ii}$ % diagonal entries not necessarily 1
end

Solving Linear Systems - LU Factorizations

Complexity

- 1 **flop** $= + / - / * / \div$.
- Consider the forward solve algorithm. For each i, the j-loop performs $2(i-1)$ flops.

$$
\begin{aligned}
\text{Total flops} &= \sum_{i=1}^{n} 2(i-1) \\
&= 2\sum_{i=1}^{n} i - \sum_{i=1}^{n} 2 \\
&= 2\frac{n(n+1)}{2} - 2n \\
&= n^2 + n - 2n \\
&= n^2 - n \\
&\in O(n^2).
\end{aligned}
$$

- flops(back-solve) $\in O(n^2)$ (Exercise).
- flops(LU factorization) $= \frac{2}{3}n^3 + O(n^2)$ (Exercise).

Solving Linear Systems - LU Factorizations

For large n, the factorization is more expensive than forward and back solving.

Solving Linear Systems - LU Factorizations

Special Linear Systems

- Exploit special structures of linear systems
- More efficient LU factorization

Solving Linear Systems - Symmetric Systems

- LDM^T factorization, variant of LU.

We do NOT assume that A is symmetric yet; the next Theorem applies whether A is symmetric or not.

Definition 2

A **principal submatrix** *is a smaller matrix constructed by deleting rows and corresponding columns.*

Solving Linear Systems - Symmetric Systems

Some examples are

$$\begin{bmatrix} \cancel{1} & \cancel{2} & \cancel{3} & \cancel{4} & \cancel{5} \\ \cancel{6} & \cancel{7} & \cancel{8} & \cancel{9} & \cancel{10} \\ \cancel{11} & \cancel{12} & \cancel{13} & \cancel{14} & \cancel{15} \\ \cancel{16} & \cancel{17} & \cancel{18} & 19 & 20 \\ \cancel{21} & \cancel{22} & \cancel{23} & 24 & 25 \end{bmatrix} \rightarrow \begin{bmatrix} 19 & 20 \\ 24 & 25 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & \cancel{2} & 3 & \cancel{4} & 5 \\ \cancel{6} & \cancel{7} & \cancel{8} & \cancel{9} & \cancel{10} \\ 11 & \cancel{12} & 13 & \cancel{14} & 15 \\ \cancel{16} & \cancel{17} & \cancel{18} & \cancel{19} & \cancel{20} \\ 21 & \cancel{22} & 23 & \cancel{24} & 25 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 5 \\ 11 & 13 & 15 \\ 21 & 23 & 25 \end{bmatrix}.$$

Solving Linear Systems - Symmetric Systems

The following results are reproduced from Chapter 3 of *Matrix Analysis and Applied Linear Algebra*, by Carl D. Meyer.

Definition 3

The **leading principal submatrices** *of A are defined to be those submatrices taken from the upper-left-hand corner of A. That is*

$$\begin{aligned}
A_1 &= \begin{bmatrix} a_{11} \end{bmatrix}, \\
A_2 &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \\
&\vdots \\
A_{n-1} &= \begin{bmatrix} a_{11} & \cdots & a_{1\ n-1} \\ \vdots & & \vdots \\ a_{n-1\ 1} & \cdots & a_{n-1\ n-1} \end{bmatrix}, \\
A_n &= A
\end{aligned}$$

Solving Linear Systems - Symmetric Systems

Theorem 4

Each of the following statements is equivalent to saying that a non-singular $n \times n$ matrix A possesses an LU-factorization.

1. *A zero pivot does not emerge during row-reduction to upper-triangular form with Type III operations.*
2. *Each leading principal submatrix A_k is non-singular.*

Solving Linear Systems - Symmetric Systems

Proof: We will prove the statement concerning the leading principal submatrices and leave the proof concerning the nonzero pivots as an exercise.

First, assume that A has an LU-factorization.

- For any $1 \leq k \leq n$, partition A as

$$\begin{aligned} A &= LU \\ &= \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} \\ &= \begin{bmatrix} L_{11}U_{11} & * \\ * & * \end{bmatrix}, \end{aligned}$$

  where L_{11} and U_{11} are both $k \times k$.
- Then $A_k = L_{11}U_{11}$ is non-singular, because both of L_{11} and U_{11} are non-singular (each is triangular, with non-zero diagonal entries).
- Since k was arbitrary, this shows that all of the leading principal submatrices of A are nonsingular.

Solving Linear Systems - Symmetric Systems

Second, assume that the leading principal submatrices of A are all non-singular.

- Let $1 \leq k \leq n$ be arbitrary.
- We will prove by induction on k that each A_k has an LU-factorization.
- Then, since $A = A_n$, it follows that A has an LU-factorization.

Solving Linear Systems - Symmetric Systems

Base ($k = 1$):

- $A_1 = \begin{bmatrix} a_{11} \end{bmatrix}$, so the assumption that A_1 is non-singular guarantees that $a_{11} \neq 0$.
- Then $A_1 = \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} a_{11} \end{bmatrix}$ is an LU-factorization of A_1.
- This completes the base case.

Solving Linear Systems - Symmetric Systems

Induction ($k > 1$):

- The induction hypothesis is that all A_ℓ, for $1 \leq \ell < k$, have LU-factorizations.
- In particular, A_{k-1} has an LU-factorization.
- Write $A_{k-1} = L_{k-1}U_{k-1}$.
- By assumption, A_{k-1} is non-singular.
- Therefore $A_{k-1}^{-1} = U_{k-1}^{-1}L_{k-1}^{-1}$.
- Define

$$\begin{aligned} c^T &= \text{the first } k-1 \text{ components of the } k^{th} \text{ row of } A_k, \\ b &= \text{the first } k-1 \text{ components of the } k^{th} \text{ column of } A_k, \\ \alpha_k &= \text{the } (k,k) \text{ entry of } A_k. \end{aligned}$$

- With this notation, we can write

$$A_k = \begin{bmatrix} A_{k-1} & b \\ c^T & \alpha_k \end{bmatrix}.$$

Solving Linear Systems - Symmetric Systems

I claim that the following is an LU-factorization of A_k:

$$\underbrace{\begin{bmatrix} L_{k-1} & 0 \\ c^T U_{k-1}^{-1} & 1 \end{bmatrix}}_{L_k} \underbrace{\begin{bmatrix} U_{k-1} & L_{k-1}^{-1} b \\ 0 & \alpha_k - c^T A_{k-1}^{-1} b \end{bmatrix}}_{U_k}$$

Solving Linear Systems - Symmetric Systems

We verify that this works in each of the 4 blocks of A_k.

- top-left $k-1 \times k-1$ block: $L_{k-1}U_{k-1} = A_{k-1}$, by the induction hypothesis.
- row k, first $k-1$ entries: $\underbrace{\begin{bmatrix} c^T U_{k-1}^{-1} & 1 \end{bmatrix}}_{1\times k} \underbrace{\begin{bmatrix} U_{k-1} \\ 0 \end{bmatrix}}_{k\times k-1} = c^T$.
- column k, first $k-1$ entries:
  $\underbrace{\begin{bmatrix} L_{k-1} & 0 \end{bmatrix}}_{k-1\times k} \underbrace{\begin{bmatrix} L_{k-1}^{-1} b \\ \alpha_k - c^T A_{k-1}^{-1} b \end{bmatrix}}_{k\times 1} = b$.
- k, k entry: $\underbrace{\begin{bmatrix} c^T U_{k-1}^{-1} & 1 \end{bmatrix}}_{1\times k} \underbrace{\begin{bmatrix} L_{k-1}^{-1} b \\ \alpha_k - c^T A_{k-1}^{-1} b \end{bmatrix}}_{k\times 1} =$
  $c^T \underbrace{U_{k-1}^{-1} L_{k-1}^{-1}}_{=A_{k-1}^{-1}} b + \alpha_k - c^T A_{k-1}^{-1} b = \alpha_k$.

Solving Linear Systems - Symmetric Systems

- Observe that
  - L_k has 1s on its diagonal, and
  - U_k has non-zeros on its diagonal. The fact that $\alpha_k - c^T A_{k-1}^{-1} b \neq 0$ follows, because A_k and L_k are both non-singular, hence $U_k = L_k^{-1} A_k$ must also be non-singular.
- This completes the induction step, and hence the proof.

$\square$

Theorem 5

If all the leading principal submatrices of A are nonsingular, then there exist unique unit lower diagonal matrices L and M, and a unique diagonal matrix D such that

$$A = LDM^T$$

Solving Linear Systems - Symmetric Systems

Proof.

- Uniquely factor $A = LU$, with L unit lower triangular.
- Define $D = diag(d_1, \ldots, d_n), d_i = u_{ii},\ 1 \leq i \leq n$.
- Note, all $d_i \neq 0$, by the hypothesis that A's leading principal submatrices are all non-singular.
- Hence D^{-1} exists (its diagonal entries are the reciprocals of D's diagonal entries).
- Let $M^T = D^{-1}U$.
- Observe that $D^{-1}U$ is upper triangular, and moreover, it has 1s on its diagonal (by the construction of D^{-1}).
- This says that M^T is unit upper triangular.
- Therefore M is unit lower triangular.
- Thus $A = LU = LD(D^{-1}U) = LDM^T$.

□

Solving Linear Systems - Symmetric Systems

Remarks:

1. LU-factorization, and hence LDM-factorization, lie in $O(n^3)$.

Theorem 6

Keep the hypotheses on A from Theorem 5. If A is symmetric, then $A = LDL^T$.

Solving Linear Systems - Symmetric Systems

Proof.

- By Theorem 5, there is a unique factorization $A = LDM^T$.
- Since A is symmetric, we have

$$
\begin{aligned}
A &= A^T \\
&= (LDM^T)^T \\
&= (M^T)^T D^T L^T \\
&\underbrace{=}_{D \text{ is diagonal, hence symmetric}} MDL^T.
\end{aligned}
$$

- By the uniqueness of the LDM-factorization, we have $M = L$.

$\square$

Solving Linear Systems - Positive Definite Systems

Definition 7

An $n \times n$ symmetric matrix A is **positive definite** *if $x^T Ax > 0$, for all non-zero $n \times 1$ matrices x.*

Roughly speaking, Definition 7 generalizes a definition for positive scalars, i.e., $a \in \mathbb{R}$ is positive if $xax > 0, \forall x \in \mathbb{R}, x \neq 0$. Consider the function $f(x) = x^T Ax$, which is quadratic and A contains the coefficients. Positive definiteness essentially asks if f is convex.

Solving Linear Systems - Positive Definite Systems

Figure 1 shows examples of f with different A matrices.

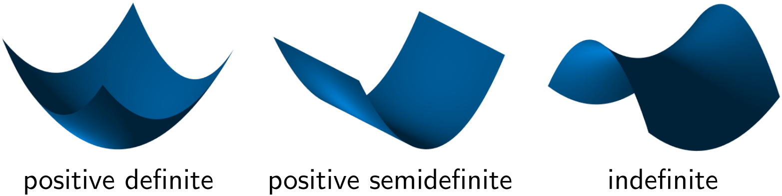

Figure: Plotting of $f = x^T A x$ as a height function with $x \in \mathbb{R}^2$ and different $A \in \mathbb{R}^{2\times 2}$.

Solving Linear Systems - Positive Definite Systems

Remarks:

1. There is not universal agreement among mathematicians that we should only ask if a matrix is positive definite if we already know that it is symmetric.
2. E.g. $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ satisfies $x^T Ax > 0$, for all $0 \neq x \in \mathbb{R}^2$, but is clearly not symmetric.
3. In this course, we will only care if a matrix is positive definite when we already know that it is symmetric.
4. This is why we make being symmetric part of the definition of being positive definite.
5. Keep in mind that the definition of positive definiteness may be different in other contexts, outside of this course.

Solving Linear Systems - Positive Definite Systems

Equivalent Characterizations of Being Positive Definite

1. All eigenvalues are strictly positive. (This makes sense because the eigenvalues of a symmetric matrix are all real.)
2. All pivots are strictly positive. (using the fact that for a symmetric matrix, the signs of the pivots are the signs of the eigenvalues.)
3. The k^{th} pivot of a matrix is

$$d_k = \frac{\det(A_k)}{\det(A_{k-I})},$$

   where A_k is the k^{th} leading principal submatrix. All the pivots will be positive if and only if $det(A_k) > 0$ for all $1 \leq k \leq n$. So, if the determinants of the leading principal submatrices are positive, then the matrix is positive definite.
4. A matrix A is positive definite if and only if it can be written as $A = R^T R$ for some possibly rectangular matrix R with independent columns.

Solving Linear Systems - Positive Definite Systems

Remarks:

1. I will not prove these equivalences in class.
2. If any student specifically asks to see the proofs, then I will type them up and post them.
3. You can safely infer from the above two comments that you are not responsible for knowing these proofs for this course.

Solving Linear Systems - Positive Definite Systems

Theorem 8

If A is positive definite, then A^{-1} exists.

A useful result for PD matrices is given in Theorem 9.

Theorem 9

If $A \in \mathbb{R}^{n \times n}$ is PD and $X \in \mathbb{R}^{n \times k}$ has rank $k \leq n$, then $B = X^T A X$ is also PD (i.e., $z^T B z > 0, \forall z \in \mathbb{R}^k, z \neq 0$).

$$\begin{bmatrix} \\ \quad A_{n \times n} \quad \\ \\ \end{bmatrix} \quad \begin{bmatrix} \\ \quad X_{n \times k} \quad \\ \\ \end{bmatrix} \quad \begin{bmatrix} \quad B_{k \times k} \quad \end{bmatrix} \quad \begin{bmatrix} z_1 \\ \vdots \\ z_k \end{bmatrix} \quad \begin{bmatrix} x_1 \\ \\ \vdots \\ \\ x_n \end{bmatrix}$$

The above diagram shows the sizes of all the matrices/vectors in Theorem 9.

Solving Linear Systems - Positive Definite Systems

Proof.

- Consider any $0 \neq z \in \mathbb{R}^k$, then

$$z^T B z = z^T X^T A X z = (Xz)^T A (Xz).$$

- Let $x = Xz$, which is a vector in $\mathbb{R}^n$.
- If $x \neq 0$ then we are finished because $(Xz)^T A (Xz) = x^T A x > 0$ since A is PD.
- When can $x = 0$? This is equivalent to asking what the null space of X is.
- Since X has rank k it is full rank.
- By the rank-nullity theorem $\dim(\text{null}(X)) = \text{nullity}(X) = 0$.
- Hence, the null space of X contains only the zero-vector.
- Thus, $x = 0$ only if $z = 0$. So $z^T B z = (Xz)^T A (Xz) > 0$ for all $z \neq 0 \Rightarrow B$ is PD.

$\square$

Solving Linear Systems - Positive Definite Systems

Corollary 10

If A is PD, then all its principal submatrices are PD. In particular, all diagonal entries are positive.

Proof.

Each diagonal entry is a principal submatrix with all other rows/columns deleted. You can design (identity-like) matrices X to "pick out" arbitrary principal submatrices using $X^T AX$ (which is PD by Theorem 9), e.g.,

$$A = \begin{bmatrix} -1 & 2 & 5 \\ 2 & 4 & -4 \\ 5 & -4 & 7 \end{bmatrix}, \qquad X = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \Rightarrow X^T AX = \begin{bmatrix} -1 & 5 \\ 5 & 7 \end{bmatrix}.$$

□

Remarks:

1. The converse of the Corollary 10 holds, because A is a principal submatrix of itself.

Solving Linear Systems - Positive Definite Systems

Corollary 11

If A is PD, so that $A = LDL^T$, then the diagonal matrix D has strictly positive entries.

Proof.

- Since L is unit lower triangular, therefore it is invertible. Hence we have

$$\begin{aligned} A &= LDL^T \\ L^{-1}AL^{-T} &= D. \end{aligned}$$

- By Theorem 9, $D = L^{-1}AL^{-T}$ is PD.
- By Corollary 10, D's diagonal entries are all positive.

$\square$