Math 170A: Introduction to Numerical Analysis: Linear Alg.

Instructor: Thang Huynh

Course Webpage: thanghuynh.org → teaching

→ This course.

* Contains / will contain: Syllabus
  Exam Schedule
  HW
  Office Hours.
  TA Information.
  etc.

* Grading scheme: HW, 2MTs, Final.

20%, 20%, 20%, 40% Final

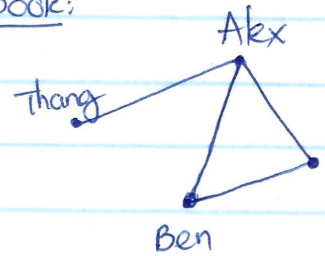
or 20% (HW), 20% (highest MT), 60% Final

(whichever is higher)

- What is numerical linear algebra (NLA)?
  Solving linear algebra problems using efficient algos. on computers.
- Why learn NLA?
  NLA is foundation of scientific computations.
  Many problems ultimately reduce to linear algebra concepts or algorithms. (e.g. Social Network, Machine learning.)

Talk about ML first. E.g. Uber self-driving cars:

Facebook:



Adjacency Matrix

$$\begin{array}{c|cccc} & \text{Thang} & \text{Alex} & \text{Jenny} & \text{Ben} \\ \hline \text{Thang} & 0 & 1 & 0 & 0 \\ \text{Alex} & 1 & 0 & 1 & 1 \\ \text{Jenny} & 0 & 1 & 0 & 1 \\ \text{Ben} & 0 & 1 & 1 & 0 \end{array}$$

②

~~Def.~~ Photos are represented by matrices.

In Machine Learning, say you have 1 million photos of animals, i.e. 1 mil very large matrices, how can you classify them? (Warning: we will not study this problem in this class.)

What's a matrix?

~~Def.~~

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

A is an $n \times m$ matrix.
n rows and m columns.
All entries belong to $\mathbb{R}$.

the space of $n \times m$ matrices is $\mathbb{R}^{n \times m}$.

E.g.

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1.5 & 2.3 \\ 0 & 3 \end{bmatrix}.$$

~~when~~ $m \times 1 \Rightarrow$ Vectors:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

- You can consider $\vec{x}$ as an ~~m×1 matrix~~
- $\vec{x}$ is size m.
- Space of m-vectors is $\mathbb{R}^m$

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

-

- Matrix-vector product $\vec{b} = A\vec{x}$ is defined as

$$b_i = \sum_{j=1}^{m} a_{ij} x_j$$

~~E.g.~~ $\cancel{[1 \quad 2 \quad 3]}$

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}}_{\vec{b}}$$

E.g.

1) $$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 1\cdot1 + 2\cdot1 + 3\cdot1 = 6$$

2) $$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4\cdot1 + 5\cdot1 + 6\cdot1 \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \end{bmatrix}.$$

<u>Goals:</u> Linear system

Given $A\vec{x} = \vec{b}$,

unknown.

~~How to~~ find $\vec{x}$ efficiently.

Another way to view matrix-vector product:

$$A = \begin{bmatrix} \boxed{\vec{a}_1} & \boxed{\vec{a}_2} & \cdots & \boxed{\vec{a}_m} \end{bmatrix}$$

$\vec{a}_j$ is the jth column of A.

④

$$\vec{b} = A\vec{x} = \sum_{j=1}^{m} x_j \vec{a}_j$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots + x_m \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix}$$

i.e. b is a linear combination of column vectors of A.

(<u>Recall</u>: If $\vec{v}_1, \ldots, \vec{v}_m$ are vectors, then the linear combination of those vectors with ~~real~~ (scalar) coefficients $\alpha_1, \ldots, \alpha_m$ is

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha \vec{v}_m .)$$

<u>E.g.</u>

$$\underset{\vec{b}}{\begin{bmatrix} 6 \\ 15 \end{bmatrix}} = \underset{x_1}{1} \cdot \underset{\vec{a}_1}{\begin{bmatrix} 1 \\ 4 \end{bmatrix}} + 1 \cdot \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 1 \cdot \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

. Two different views of matrix-vector products.

1) $b_i = \sum_{j=1}^{m} a_{ij} x_j$: A acts on x to produce b; scalar operations.

2) $b = \sum_{j=1}^{m} x_j \vec{a}_j$: x acts on A to produce b; Vector operations.

~~Pseudo~~

Pseudo-code for $\vec{b} = A\vec{x}$

```
for i = 1 to n.
    b_i ← 0
    for j = 1 to m
        b_i ← b_i + a_ij x_j
```

E.g. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

Pseudo-code are not real code, i.e. they do not run on any computer. They are human readable, but they should be straightforward to convert into real computer codes in any programming language (e.g. MATLAB, Python)

* Flop counts:

It is important to evaluate the efficiency of algos. But how?

→ can implement different algos. and do head-to-head comparison, ~~but implementation details can~~

→ can estimate cost of all operations, but it is very tedious.

→ Relative simple and effective approach is to estimate amount of floating-point operations, or "flops".

* Idealization:

– Count each operation $+$, $-$, $*$, $/$, and $\sqrt{\ }$, as one flop

e.g. Matrix-vector product requires about $2mn$ flops, because $b_i \leftarrow b_i + \underbrace{a_{ij}\, x_j}_{a_{ij} * x_j}$ requires two flops.

6

Suppose $m = n$, it takes quadratic time in n, i.e. $O(n^2)$.

* <u>Matrix – Matrix Multiplication:</u>

If A is $n \times m$ and X is $m \times p$, then $B = AX$ is $n \times p$, with entries defined by

$$b_{ij} = \sum_{k=1}^{m} a_{ik} x_{kj}$$

Written in columns, we have

$$\vec{b}_j = A\vec{x}_j = \sum_{k=1}^{m} x_{kj} \vec{a}_k$$

In other words, jth column of B is A times jth column of X.

- <u>Pseudo-code:</u> $B = AX$.

```
for i = 1 to n.
    for j = 1 to p
        b_ij <- 0
        for k = 1 to m
            b_ij <- b_ij + a_ik x_kj .
```

How many flops? ($2mnp$.).

If $m = n = p \rightarrow$ takes $O(n^3)$ time in n.

7

* Block Matrices:

Matrix can be partitioned into blocks.

~~E.g.~~ $A \in \mathbb{R}^{n\times m}$ $\qquad X^{m\times p}$

$$A = \begin{array}{c c} \begin{array}{cc} m_1 & m_2 \end{array} & \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] & \begin{array}{c} n_1 \\ n_2 \end{array} \end{array} \qquad X = \begin{array}{c c} \begin{array}{cc} p_1 & p_2 \end{array} & \\ \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} & \begin{array}{c} m_1 \\ m_2 \end{array} \end{array}.$$

Then $B = AX$ can be written as

$$\begin{array}{c c} \begin{array}{cc} p_1 & p_2 \end{array} & \\ \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} & \begin{array}{c} n_1 \\ n_2 \end{array} \end{array} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}.$$

where $B_{ij} = A_{i1} X_{1j} + A_{i2} X_{2j} \qquad i,j = 1,2.$

- More generally, the matrix can be decomposed into more than two blocks rows/columns.

E.g. $A = \left[\begin{array}{c|cc} 1 & 3 & 2 \\ 2 & 1 & 1 \\ \hline -1 & 0 & 1 \end{array}\right] \qquad X = \left[\begin{array}{c|cc} 1 & 0 & 1 \\ \hline 2 & 1 & 1 \\ -1 & 2 & 0 \end{array}\right]$

$$B = \left[\begin{array}{c|cc} 5 & 7 & 4 \\ 3 & 3 & 3 \\ \hline -2 & 2 & -1 \end{array}\right]$$

Show that $A_{11} X_{11} + A_{12} X_{21} = B_{11}$.

- Why do we care?

E.g. $A = \left[\begin{array}{c|cc} 2 & 0 & 0 \\ \hline 0 & 1 & 3 \end{array}\right] \qquad X = \left[\begin{array}{c|c} 1 & 2 \\ \hline 3 & 4 \\ 5 & 6 \end{array}\right]$

1.2 System of Linear Equations.

In scientific computations, one of the most common problem is to solve

(*) $$A\vec{x} = \vec{b}$$

where A is $n\times n$. That is, find x such that ~~Ax~~ $=b$.

. This is the inverse problem of matrix-vector multiplication.

(*) has a unique solution if and only if A is nonsingular

Def: Given a nonsingular matrix A, its inverse is written by A^{-1}, and $AA^{-1} = A^{-1}A = I$.

~~If A is nonsingular,~~

$$\cancel{A^{-1}A\vec{x} = A^{-1}\vec{b}}$$

$$\boxed{\cancel{\vec{x} = A^{-1}\vec{b}}}$$

If A is nonsingular, then

A^{-1} exists.

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

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

First calculate A^{-1}

then multiply A^{-1} by $\vec{b}$ to obtain $\vec{x}$.

⇒ work in theory

but is a bad idea to practice.

⇒ As we shall see, it is generally more efficient to solve $A\vec{x} = \vec{b}$ directly, without calculating A^{-1}.

9

Thm: (Existence and Uniqueness of Inverse)

Let A be an $n \times n$ square matrix. The following conditions are equivalent:

(i) A has an inverse A^{-1}.

(ii) there is no nonzero $\vec{y}$ such that $A\vec{y} = 0$

(iii) the columns of A are linearly independent

(iv) the rows of A are linearly independent.

(v) $\det(A) \neq 0$

(vi) given any vector $\vec{b}$, there is exactly one vector $\vec{x}$ such that $A\vec{x} = \vec{b}$.

In addition, 0 is not an eigenvalue of A, and 0 is not a singular value of A.

E.g. Find the intersection ~~of~~ point of the following two lines.

$$2x_1 + x_2 = 7$$

and

$$x_1 - 3x_2 = -7$$

Soln:

$$\underbrace{\begin{bmatrix} 2 & 1 \\ 1 & -3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 7 \\ -7 \end{bmatrix}}_{\vec{b}}.$$

$\det(A) = 2(-3) - 1 = -7. \neq 0. \Rightarrow$ unique sol.

$$A^{-1} = \frac{1}{-7}\begin{bmatrix} -3 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\vec{x} = A^{-1}\vec{b} = -\frac{1}{7}\begin{bmatrix} -3 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 7 \\ -7 \end{bmatrix} = \begin{bmatrix} -3 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3-1 \\ 1+2 \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

On Monday, we studied system of linear eqs. That is

Find $\vec{x}$

such that $A\vec{x} = \vec{b}$.

We know that if A is nonsingular,

$$\vec{x} = A^{-1}\vec{b}.$$

But it's not an easy task to compute A^{-1}.

~~How~~ If A has some "nice" structure?, can we ~~do~~ it better?

E.g. Find $\vec{x}$ such that

$$\underbrace{\begin{bmatrix} 5 & 0 & 0 \\ 2 & -4 & 0 \\ 1 & 2 & 3 \end{bmatrix}}_{\text{lower-triangular matrix}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 15 \\ -2 \\ 10 \end{bmatrix}$$

Sol: We first see that

$$5x_1 = 15$$

$$x_1 = 3.$$

Since

$$2x_1 - 4x_2 = -2,$$

$$-4x_2 = -2 - 2x_1$$

$$x_2 = \frac{-2 - 2x_1}{-4}$$

$$= \frac{+1 + 1 \cdot x_1}{2}$$

$$= \frac{1+3}{2}$$

$$= 2.$$

Similarly,

$$x_1 + 2x_2 + 3x_3 = 10$$

$$x_3 = \frac{10 - x_1 - 2x_2}{3}$$

$$= \frac{10 - 3 - 2(2)}{3}$$

$$= 1$$

$$\Rightarrow \vec{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

(Try to solve it using A^{-1} and compare to the above way!)

* <u>Triangular Systems</u>:

A matrix $G = (g_{ij})$ is <u>lower triangular</u> if $g_{ij} = 0$ whenever $i < j$.

$$G = \begin{bmatrix} g_{11} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & 0 & \cdots & 0 \\ g_{31} & g_{32} & g_{33} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ g_{n1} & g_{n2} & g_{n3} & \cdots & g_{nn} \end{bmatrix}$$

- Similarly, an <u>upper triangular</u> matrix is one for which $g_{ij} = 0$ whenever $i > j$.
- A triangular matrix is one that is either upper or lower triangular.
- A triangular matrix $G \in \mathbb{R}^{n \times n}$ is nonsingular if and only if $g_{ii} \neq 0$ for $i = 1, \ldots, n$.

12

* Lower – Triangular Systems.

Consider system

$$G\vec{y} = \vec{b}$$

unknown.

where G is nonsingular, lower – triangular matrix. We can solve the system by

$$y_1 = b_1 / g_{11}$$

$$y_2 = (b_2 - g_{21} y_1)/g_{22}$$

$$\vdots$$

$$y_i = (b_i - g_{i1}y_1 - g_{i2}y_2 - \dots - g_{i,i-1}y_{i-1})/g_{ii}$$

$$= \left(b_i - \sum_{j=1}^{i-1} g_{ij} y_j\right) / g_{ii}.$$

(row-oriented)

• Forward Substitution:

Pseudo – code forward substitution

```
line 1   for i = 1 to n
     2      for j = 1 to i-1
     3         b_i <- b_i - g_ij b_j
     4      ~~b_i <- b_i/g_ii~~
            if g_ii = 0, set error flag, exit
     5      b_i <- b_i/g_ii
```

• This algorithm is row-oriented, as it access G by rows.

• It may raise an exception if $g_{ii} = 0$

• The number of operations: no flops in lines 1, 2, & 4.

$$\cancel{\sum_{i=1}^{n}\sum_{j=1}^{i-1} 2 = 2\sum_{i=1}^{n}(i-1) = n(n-1) \approx n^2}.$$

	Line 3 # of flops	line 5 # of flops
$i=1$	0	1
$i=2$	2	1
$\vdots$		
$i=n$	$2(n-1)$	1.

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

$$= n(n-1) + n$$

$$\boxed{= n^2}.$$

(13)

* The column-oriented form of the forward substitution:

$$\begin{bmatrix} 5 & 0 & 0 \\ 2 & -4 & 0 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 15 \\ -2 \\ 10 \end{bmatrix}$$

$$\begin{bmatrix} 5 & \vec{0} \\ \hat{h} & \hat{G} \end{bmatrix} \begin{bmatrix} y_1 \\ \hat{y} \end{bmatrix} = \begin{bmatrix} 15 \\ \hat{b} \end{bmatrix}.$$

$$5y_1 = 15$$
$$\hat{h} y_1 + \hat{G}\hat{y} = \hat{b}. \rightsquigarrow \begin{bmatrix} 2 \\ 1 \end{bmatrix} y_1 + \begin{bmatrix} -4 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 10 \end{bmatrix}$$

$$y_1 = 15/5 = 3$$
$$\hat{G}\hat{y} = \hat{b} - 3\hat{h}.$$
$$\boxed{\hat{G}\hat{y} = \tilde{b}}$$

Then solve $\hat{G}\hat{y} = \tilde{b}$ for $\hat{y}$.

for $j = 1$ to n
  if $g_{jj} = 0$, set error flag, exit.
  $b_j \leftarrow b_j / g_{jj}$
  for $i = j+1, \ldots$ to n
    $b_i \leftarrow b_i - g_{ij} b_j$

* Upper triangular linear systems:

$$U\vec{x} = \vec{b}$$

$$\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & & \\ 0 & 0 & \cdots\ 0 & u_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_n \end{bmatrix}$$ $\Rightarrow$ last row:

$$u_{nn} x_n = b_n.$$

$\Rightarrow$ solve backward.

$$x_n = b_n / u_{nn}$$

$$x_{n-1} = (b_{n-1} - u_{n-1,n}\, x_n) / u_{n-1,n-1}$$

$$\vdots$$

$$x_i = (b_i - u_{i,n} x_n - u_{i,n-1} x_{n-1} - \cdots - u_{i,i+1} x_{i+1}) / u_{ii}$$

$$= \left(b_i - \sum_{j=i+1}^{n} u_{ij}\, x_j\right) / u_{ii}.$$

* Section 1.4: Positive definite systems; Cholesky Decomposition.

~~Def: A real symmetric matrix A is positi~~

Def: An $n \times n$ matrix A is symmetric if $A^T = A$.

E.g. $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ symmetric but $\underset{A}{\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}}$ not since $\underset{A^T}{\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}}$.

Def: A real symmetric matrix A is positive definitive if $\underset{1\times n}{x^T}\underset{n\times n}{A}\underset{n\times 1}{x} > 0$ for all vectors $x \neq 0$.

Note: if A is a real symmetric matrix and $x^T A x \geq 0$ for all vectors $x \neq 0$, then A is called positive semi-definite.

E.g. $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Is A positive definite?

$$
\begin{aligned}
x^T A x &= [x_1 \quad x_2]\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [2x_1 + x_2 \quad x_1 + 2x_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= (2x_1 + x_2)x_1 + (x_1 + 2x_2)x_2 \\
&= 2x_1^2 + x_1x_2 + x_1x_2 + 2x_2^2 \\
&= 2x_1^2 + 2x_1x_2 + 2x_2^2 \\
&= x_1^2 + 2x_1x_2 + x_2^2 + x_1^2 + x_2^2 \\
&= (x_1 + x_2)^2 + x_1^2 + x_2^2 \\
&> 0 \quad \text{since } x \neq 0.
\end{aligned}
$$

$\Rightarrow$ Yes.

Thm: If A is positive definite, then A is nonsingular.

Thm: Let M be any $m \times n$ matrix. Then $A = M^T M$ is psd.

Pf: For any $x \neq 0$

$$
\begin{aligned}
x^T A x &= x^T M^T M x \\
&= (Mx)^T Mx. \\
\text{let } v = Mx \qquad &= v^T v \\
&= \|v\|^2 \\
&\geq 0.
\end{aligned}
$$

$\Rightarrow A$ is psd.

Corollary: Let M be any $n \times n$ nonsingular matrix ~~any $m \times n$ real matrix with rank equal to n~~. Then $A = M^T M$ is p.d.

E.g. Let $M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. Since $\det(M) = 4 - 6 = -2 \neq 0$, M is nonsingular.

Then $A = M^T M = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 14 \\ 14 & 20 \end{bmatrix}$ is p.d.

* Cholesky Decomposition:

Thm: If A is symmetric positive definite, then there is a unique factorization of A such that

$$A = R^T R = L L^T$$

where R is upper triangular, ~~and~~ all its diagonal entries are positive, and $L = R^T$ is a lower triangular.

E.g. Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}$. Then

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}}_{R^T} \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{R}.$$

Why is it helpful?

Suppose A is p.d. then $A = R^T R$.

Solve $A\vec{x} = \vec{b}$ for $\vec{x}$.

$R^T R \vec{x} = \vec{b}$.

Let $\vec{y} = R\vec{x}$.

Solve $R^T \vec{y} = \vec{b}$ for $\vec{y}$.

Then solve $R\vec{x} = \vec{y}$ for $\vec{x}$.

$$\cancel{A = \begin{bmatrix} a_{11} & b^T \\ b & K \end{bmatrix}}$$

(SPD)

* Properties of Symmetric Positive-Definite Matrices.

1) If A is SPD, then A is nonsingular.

2) Let X be any $n\times m$ matrix with full rank and $n\geq m$. Then

- X^TX is ~~SPD and~~ symmetric positive definite, and
- XX^T is symmetric positive semidefinite.

3) If A is $n\times n$ SPD and $X\in\mathbb{R}^{n\times m}$ has full rank and $n\geq m$, then X^TAX is SPD.

4) Any principal submatrix (picking some rows and corresponding columns) of A is SPD and $a_{ii}>0$.

* Cholesky Factorization:

If A is symmetric positive definite, then there is a factorization of A such that

$$A = R^TR$$

where R is upper triangular, and all its diagonal entries are positive.

Pf: By induction. $A=[a_{11}] \Rightarrow R=[\sqrt{a_{11}}]$ ✓. Suppose theorem is true for all $(n-1)\times(n-1)$ SPD matrices.

$$A=\begin{bmatrix} a_{11} & b^T \\ b & K\end{bmatrix}$$

check

$$=\begin{bmatrix} r_{11} & 0 \\ b/r_{11} & I\end{bmatrix}\begin{bmatrix} r_{11} & b^T/r_{11} \\ 0 & K-bb^T/r_{11}\end{bmatrix}, \text{ where } r_{11}=\sqrt{a_{11}}$$

$$=\begin{bmatrix} r_{11} & 0 \\ b/r_{11} & I\end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & K-bb^T/r_{11}\end{bmatrix}\begin{bmatrix} r_{11} & b^T/r_{11} \\ 0 & I\end{bmatrix}$$

$$\Rightarrow A = R_1^T \quad A_1 \quad R_1.$$

- $K-bb^T/r_{11}$ is principal submatrix of SPD $A_1=R_1^{-T}AR_1^{-1}$

and therefore is SPD, with positive diagonal entries.

- Apply recursively, to obtain

$$A = (R_1^T R_2^T \cdots R_n^T)(R_n \cdots R_2 R_1) = R^T R, \quad r_{jj} > 0.$$

which is known as Cholesky factorization.

$\Rightarrow$ • $K - bb^T/r_{11}$ is a $(n-1)\times(n-1)$ SPD.

$\Rightarrow$ by inductive hypothesis, There exists $\tilde{R}$ s.t.

$$K - bb^T/r_{11} = \tilde{R}^T \tilde{R}.$$

$$\Rightarrow A = \begin{bmatrix} r_{11} & 0 \\ b/r_{11} & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & \tilde{R}^T\tilde{R} \end{bmatrix}\begin{bmatrix} r_{11} & b^T/r_{11} \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} r_{11} & 0 \\ b/r_{11} & I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & \tilde{R}^T \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & \tilde{R} \end{bmatrix}\begin{bmatrix} r_{11} & b^T/r_{11} \\ 0 & I \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} r_{11} & 0 \\ b/r_{11} & \tilde{R}^T \end{bmatrix}}_{R^T}\underbrace{\begin{bmatrix} r_{11} & b^T/r_{11} \\ 0 & \tilde{R} \end{bmatrix}}_{R}$$

$$= R^T R.$$

- Every SPD matrix has a unique Cholesky factorization
  - $\rightarrow$ Exists because algorithms for Cholesky factorization always works for SPD matrices.
  - $\rightarrow$ Unique because once $r_{11} = \sqrt{a_{11}}$ is determined at each step, entire column b/r_{11} is determined.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} r_{11} & 0 & 0 & \cdots & 0 \\ r_{12} & r_{22} & 0 & \cdots & 0 \\ r_{13} & r_{23} & r_{33} & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ r_{1n} & r_{2n} & r_{3n} & \cdots & r_{nn} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

The element a_{ij} is the inner product of the ith row of R^T with the jth column of R.

1st row:

$$a_{1j} = r_{11} r_{1j} + 0 + 0 + \dots + 0.$$

For $j = 1$, $\quad a_{11} = r_{11}^2 \Rightarrow r_{11} = +\sqrt{a_{11}}$.

$$\Rightarrow \quad r_{1j} = \frac{a_{1j}}{r_{11}}.$$

2nd row:

$$a_{2j} = r_{12} r_{1j} + r_{22} r_{2j}$$

$j = 2$, $\quad a_{22} = r_{12}^2 + r_{22}^2 \Rightarrow r_{22} = \sqrt{a_{22} - r_{12}^2}$

$$r_{2j} = \frac{a_{2j} - r_{12} r_{1j}}{r_{22}}.$$

Now suppose that we already have the first $i-1$ rows.

$$a_{ij} = r_{1i} r_{1j} + r_{2i} r_{2j} + \dots + r_{i-1,i} r_{i-1,j} + r_{ii} r_{ij}.$$

Take $j = i$,

$$a_{ii} = r_{1i}^2 + r_{2i}^2 + \dots + r_{i-1,i}^2 + r_{ii}^2.$$

$$r_{ii} = +\sqrt{a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2}$$

and $\quad r_{ij} = \left(a_{ij} - \sum_{k=1}^{i-1} r_{ki} r_{kj} \right) / r_{ii}$.

E.g. $A = \begin{bmatrix} 4 & -2 & 4 \\ -2 & 10 & -2 \\ 4 & -2 & 8 \end{bmatrix}$

$r_{11} = \sqrt{a_{11}} = \sqrt{4} = 2.$

$a_{1j} = r_{11} r_{1j} \Rightarrow \quad r_{12} = \dfrac{a_{12}}{r_{11}} = -\dfrac{2}{2} = -1.$

$r_{13} = \dfrac{a_{13}}{r_{11}} = \dfrac{4}{2} = 2.$

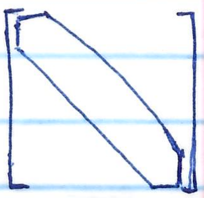
$r_{22} = \sqrt{a_{22} - r_{12}^2} = \sqrt{10 - (-1)^2} = \sqrt{9} = 3$

$r_{23} = \dfrac{a_{23} - r_{13} r_{12}}{r_{22}} = \dfrac{-2 - 2(-1)}{3} = 0.$

$r_{33} = \sqrt{a_{33} - r_{13}^2 - r_{23}^2} = \sqrt{8 - 2^2 - 0^2} = \sqrt{4} = 2.$

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

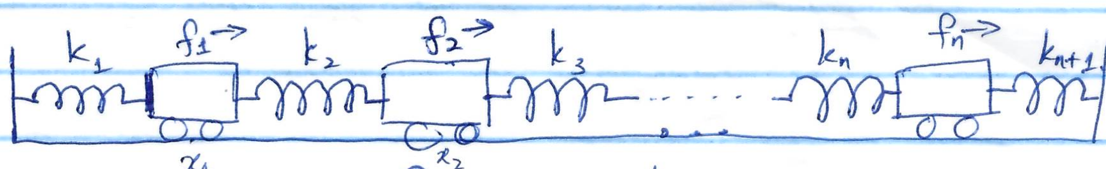
Banded Positive Definite Systems:



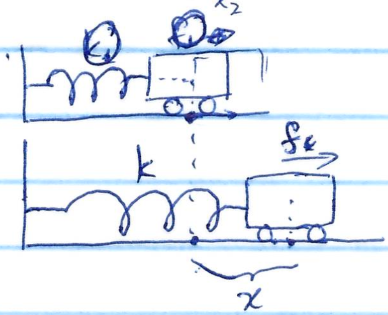
- A matrix A is banded if there is a narrow band around the main diagonal such that all of the entries of A outside of the band are zero.
- If A is $n \times n$, and there is an $s \ll n$ such that $a_{ij} = 0$ wherever $|i-j| > s$, then we say that A is banded with bandwidth $2s+1$.
- For symmetric matrices, only half of band is stored. We say that A has semi-bandwidth s.

Thm: Let A be a banded, symmetric positive def. matrix with semi-bandwidth s. Then its Cholesky factor R also has semi-bandwidth s.

E.g.

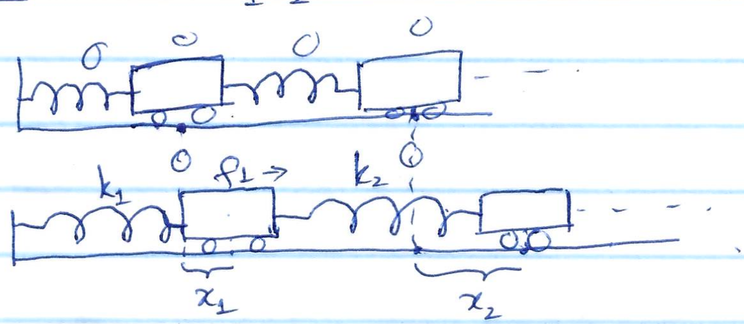


Let's consider



equilibrium. $-k(x-0) + f = 0$.

1st cart: $\cancel{k_1 x_1} +$



$$-k_1(x_1) + k_2(x_2 - x_1) + f_1 = 0.$$
$$-(k_1 + k_2)x_1 + k_2 x_2 + f_1 = 0$$
$$(k_1 + k_2)x_1 + -k_2 x_2 = f_1.$$

2nd cart:

$$-k_2(x_2 - x_1) + k_3(x_3 - x_2) + f_2 = 0.$$
$$k_3 x_3 - (k_2 + k_3)x_2 + k_2 x_1 + f_2 = 0.$$
$$\cancel{k_2 x_3}$$
$$-k_2 x_1 + (k_2 + k_3)x_2 - k_3 x_3 = +f_2.$$

ith cart:

$$-k_i x_{i-1} + (k_i + k_{i+1})x_i - k_{i+1}x_{i+1} = f_i.$$

nth cart.

$$-k_n(x_n - x_{n-1}) + k_{n+1}(-x_n) + f_n = 0.$$
$$+k_n x_{n-1} - (k_n + k_{n+1})x_n + f_n = 0.$$
$$-k_n x_{n-1} + (k_n + k_{n+1})x_n = f_n.$$

$$\begin{bmatrix} (k_1+k_2) & -k_2 & 0 & \cdots & 0 & 0 \\ -k_2 & k_2+k_3 & -k_3 & 0 \cdots & 0 & 0 \\ 0 & -k_3 & k_3+k_4 & & & \\ 0 & 0 & \vdots & & & \\ 0 & 0 & 0 & & -k_n & k_n+k_{n+1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_n \end{bmatrix}$$

E.g. Suppose $f(x)$ is given for $0 \leq x \leq 1$.
Find a function $u(x)$ such that

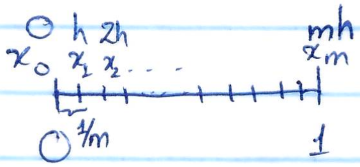
$$-u''(x) = f(x) \qquad 0 < x < 1$$

and $u(0) = T_0$, $u(1) = T_1$.

Sol:

$$u'(x) \approx \frac{u(x+h) - u(x)}{h} \qquad h \text{ very small}$$

$$u''(x) \simeq \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}$$



m very large.

$$x_i = ih \quad \text{for } i = 0, 1, 2, \ldots, m.$$

We will approximate the solution of (+) on $x_0, \ldots, x_m$.

$$-u''(x_i) = f(x_i)$$

Since

$$u''(x_i) \approx \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2}$$

$$\Rightarrow \quad \frac{-u(x_{i+1}) + 2u(x_i) - u(x_{i-1})}{h^2} \approx f(x_i).$$

Denote $u_{i+1} = u(x_{i+1})$.

$$\Rightarrow \quad -u_{i+1} + 2u_i - u_{i-1} = h^2 f(x_i), \quad i = 1, \ldots, m-1.$$

$$u_0 = T_0 \quad \text{and} \quad u_m = T_1 .$$

$\Rightarrow$ a system of $m-1$ linear equations for $m-1$ unknowns $u_1, \ldots, u_{m-1}$.

$$Au = b$$

where

~~A=~~

$$A = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & 1 \\ & & & & -1 & 2 \end{bmatrix}$$

and

$$b = \begin{bmatrix} h^2 f(x_1) + T_0 \\ h^2 f(x_2) \\ h^2 f(x_3) \\ \vdots \\ h^2 f(x_{m-2}) \\ h^2 f(x_{m-1}) + T \end{bmatrix}$$

* ~~Jacobi Method~~:

E.g.

$$\begin{aligned} 5x_1 - x_2 + 2x_3 &= 12 \\ 3x_1 + 8x_2 - 2x_3 &= -25 \\ x_1 + x_2 + 4x_3 &= 6. \end{aligned} \Rightarrow \begin{aligned} x_1 &= \frac{12 + x_2 - 2x_3}{5} \\ x_2 &= \frac{-25 - 3x_1 + 2x_3}{8} \\ x_3 &= \frac{6 - x_1 - x_2}{4} \end{aligned}$$

Iteration

	0	1	2
x_1	0	$x_1^{(1)} = \frac{12+0-2(0)}{5} = \frac{12}{5} = 2.4.$	$x_1^{(2)} = \frac{12 + (-3.125) - 2(1.5)}{5} = 1.1750$
x_2	0	$x_2^{(1)} = -25/8 = -3.125$	$x_2^{(2)} = \frac{-25 - 3(2.4) + 2(1.5)}{8} = -3.65$
x_2	0	$x_3^{(2)} = \frac{6}{4} = 1.5$	$x_3^{(2)} = \frac{6 - 2.4 + 3.125}{4} = 1.68125$

$$\vec{x}^{(0)} \longrightarrow \vec{x}^{(1)} \longrightarrow \vec{x}^{(2)}.$$

True solution is $\vec{x}^* = (1, -3, 2)$

My hope is that after many iterations, $\vec{x}^{(k)}$ is very close to (or with luck equal to) $\vec{x}^*$.

The above method is called Jacobi Method.

$\Rightarrow$ General case.

Let $A \in \mathbb{R}^{n \times n}$ be any nonsingular matrix. with $a_{ii} \neq 0$ for $i = 1, \dots, n$. Let $\vec{b} \in \mathbb{R}^n$.

Goal: Solve $A\vec{x} = \vec{b}$ for $\vec{x}$.

(27)

What's an iterative method?

- require an initial guess $\vec{x}^{(0)}$ approximating the true solution.
- Once we have $\vec{x}^{(0)}$, use it to generate $\underbrace{\vec{x}^{(1)}}_{\text{new guess}}$.
- Then use $\vec{x}^{(1)}$ to generate $\vec{x}^{(2)}$.
- and so on.

$\Rightarrow$ we generate a sequence of iterate $(\vec{x}^{(k)})$ which we hope converges to the true solution $\vec{x}^{*}$.

$\Rightarrow$ But we will not iterate forever.
Once $\vec{x}^{(k)}$ is sufficiently close to the solution (e.g. by $\|b - A\vec{x}^{(k)}\|$), we stop and accept $\vec{x}^{(k)}$ as an approximation to the solution.

* Jacobi's Method:

The ith equation in $A\vec{x} = \vec{b}$ is

$$\sum_{j=1}^{n} a_{ij} x_j = b_i .$$

$$a_{i1} x_1 + \dots + a_{i,i-1} x_{i-1} + \boxed{a_{i,i} x_i} + a_{i,i+1} x_{i+1} + \dots + a_{in} x_n = b_i$$

$$\Rightarrow \boxed{x_i = \frac{b_i - \sum_{j \neq i} a_{ij} x_j}{a_{ii}}}$$

28

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

for $i = 1, \dots, n.$

$$\vec{x}^{(k+1)} = D^{-1}\left(b + (D - A)\vec{x}^{(k)}\right).$$

* Gauss – Seidel Method.

$$\sum_{i=1}^{n} a_{i,j}\, x_j = b_i .$$

use the ith equation to modify the ith unknown, but now we consider doing the process sequentially.

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

E.g.

$$\begin{aligned} 5x_1 - x_2 + 2x_3 &= 12 \\ 3x_1 + 8x_2 - 2x_3 &= -25 \\ x_1 + x_2 + 4x_3 &= 6. \end{aligned} \Rightarrow \quad x_1 = \frac{12 + x_1 - 2x_3}{5}, \quad x_2 = \frac{-25 - 3x_1 + 2x_3}{8}, \quad x_3 = \frac{6 - x_1 - x_2}{4}$$

	0	1	2
x_1	0	$x_1^{(1)} = 2.4$	$x_1^{(2)} = 1.175$
x_2	0	$x_2^{(1)} = -3.125$	$x_2^{(2)} = \dfrac{-25 - 3(1.175) + 2(1.5)}{8} = -3.1906$
x_3	0	$x_3^{(1)} = 1.5$	$x_3^{(2)} = \dfrac{6 - 1.175 + 3.1906}{4} = 2.0039.$