Lecture 12 2.25.2019

§2-4 Symmetric positive definite (SPD) matrix and Cholesky factorization

Def: $A \in \mathbb{R}^{n\times n}$. A is symmetric if $A = A^T$. A is positive definite if $\underbrace{\underline{x}^T}_{1\times n} \underbrace{A}_{n\times n} \underbrace{\underline{x}}_{n\times 1} > 0$ for any nonzero $\underline{x} \in \mathbb{R}^n$

Theorem: Let $A \in \mathbb{R}^{n\times n}$ be symmetric. A is positive definite if and only if $\Longleftrightarrow$ all eigenvalues of A are positive.

$\Rightarrow$ A is P.D. $\Leftrightarrow$ eigenvalues are positive

$$\underline{x}^T A \underline{x} > 0 \quad \underline{x} \neq 0$$

idea: take $\underline{x}$ to be an eigenvector of an eigenvalue λ

$\Downarrow$

$\lambda > 0$

(then $\underline{x}^T A \underline{x} = \underline{x}^T \lambda \underline{x}$
$= \lambda \underline{x}^T \underline{x}$
$= \lambda \|\underline{x}\|_2^2 > 0$
$\Rightarrow \lambda > 0$)

Theorem 2 (negative results)

Assume $A \in \mathbb{R}^{n \times n}$ is symmetric.

1) A is not positive definite if some diagonal entry is negative or zero

2) A is not positive definite if the largest entry of A, in absolute value, is off the diagonal.

3) A is not positive definite if $\det(A) \leq 0$

Sketch the proof:

1) Take $\underline{x} = \underline{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow$ ith $\neq \underline{0}$

$$\underline{e}_i^T A \underline{e}_i = a_{ii}$$

2) Hint: try $\underline{x} = \underline{e}_i + \underline{e}_j$

then try $\underline{x} = \underline{e}_i - \underline{e}_j$

in $\underline{x}^T A \underline{x} > 0$

$$\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix}$$

$$\Rightarrow |a_{ij}| < \frac{a_{ii} + a_{jj}}{2} \quad \forall i, j \quad i \neq j$$

3) $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$

due to: $\det(\lambda I - A) = 0 \leftarrow$ to get eigenvalue

$$= (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

then take $\lambda = 0$

Examples: Tell whether the following symmetric matrices are positive definite.

$$A_1 = \begin{pmatrix} 4 & 1 \\ 1 & -4 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & 4 \\ 4 & 2 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 1 & 4 \\ 4 & 4 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}$$

$$A_5 = \begin{pmatrix} 2 & 4 \\ 4 & 5 \end{pmatrix}$$

A_1, A_2, A_3 : not p.D.

A_4: let $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq \underline{0}$

$$\underline{x}^T A_4 \underline{x} = (x_1, x_2) \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (x_1, x_2) \begin{pmatrix} 2(x_1 + x_2) \\ 2x_1 + 5x_2 \end{pmatrix}$$

$$= 2x_1^2 + 4x_1x_2 + \underbrace{5x_2^2}_{2x_2^2 + 3x_2^2}$$

$$= 2(x_1 + x_2)^2 + 3x_2^2 \geq 0$$

"=" will happen when $\begin{cases} x_1 + x_2 = 0 \\ x_2 = 0 \end{cases}$

$\Rightarrow x_1 = 0 = x_2$

$\Rightarrow \underline{x}^T A_4 \underline{x} > 0 \quad (\underline{x} \neq \underline{0})$

A_5 : check eigenvalues.

$$\det(\lambda I - A_5) = 0$$

$$(\lambda-2)(\lambda-5)-16=0$$

$$\lambda^2 - 7\lambda - 6 = 0$$

$$\lambda_1 \lambda_2 \overset{?}{=} -6 < 0$$, one is negative

$$(\lambda-\lambda_1)(\lambda-\lambda_2)$$ $\Rightarrow$ A_5 is not P.D.

Theorem: let $A \in \mathbb{R}^{n \times n}$ be SPD, then it always has Cholesky factorization

$$A = R^T R.$$ R is upper-triangular

and $r_{ii} > 0 \quad i=1,2,\dots n.$

Recall: cost of GE / LU factorization $\frac{2}{3}n^3$

Cost of Cholesky factorization: $\frac{1}{3}n^3$

(half of LU/GE) due to symmetry

Given $A \in \mathbb{R}^{n \times n}$, it is SPD.
Given $\underline{b} \in \mathbb{R}^n$. to solve $A\underline{x} = \underline{b}$

Step 1: find R such that

$$A = R^T R \quad \cdots \quad \frac{1}{3} n^3$$

Step 2: solve $R^T \underline{y} = \underline{b}$

solve $\underline{y}$ $\cdots$ n^2

Step 3: Solve $\underline{x}$ from

$$R \underline{x} = \underline{y} \quad \cdots \quad n^2$$

How to find Cholesky factorization

Example: $A = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix}$

$\cdots$ SPD ✓

$$A = \begin{pmatrix} 2 & 2 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} r_{11} & 0 \\ r_{12} & r_{22} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$$

$$= \begin{pmatrix} r_{11}^2 & r_{12} r_{11} \\ r_{12} r_{11} & r_{12}^2 + r_{22}^2 \end{pmatrix}$$

Compare entry by entry (1st column, then 2nd column. use symmetry so not all entries need to be examined)

$$2 = r_{11}^2 \quad (r_{11} > 0)$$

$$\Rightarrow r_{11} = \sqrt{2}$$

$$2 = r_{12} \underbrace{r_{11}}_{= \sqrt{2}} \Rightarrow r_{12} = \sqrt{2}$$

$$5 = \underbrace{r_{12}^2}_{= (\sqrt{2})^2} + r_{22}^2 \Rightarrow r_{22} = \sqrt{3}$$

$$r_{22} > 0$$

$$R = \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ & \sqrt{3} \end{pmatrix}$$

Remark: the procedure in the example can be extended to a SPD matrix $A \in \mathbb{R}^{n \times n}$ (any n).

§ 2-5 Nonlinear systems of equations.

So far we know how to solve

$f(x) = 0$ scalar (linear / nonlinear)

$A\underline{x} = \underline{b}$ system (linear)

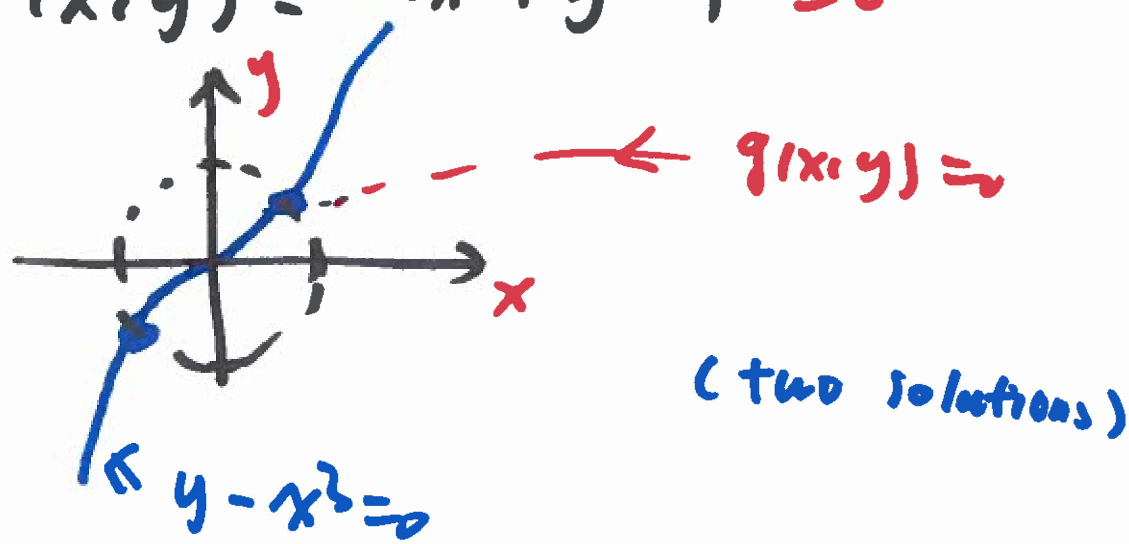
we now want to consider

$$\begin{cases} f_1(x_1, x_2, \cdots x_n) = 0 \\ f_2(x_1, x_2, \cdots x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \cdots x_n) = 0 \end{cases}$$

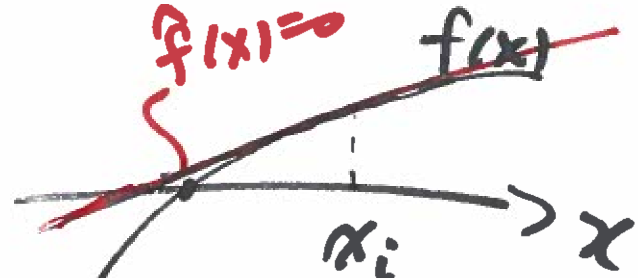
and solve for $x_1, x_2, \cdots x_n$

Example $n=2$

$$\begin{cases} f(x,y) = y - x^3 = 0 \\ g(x,y) = x^2 + y^2 - 1 = 0 \end{cases}$$



Newton's method:



Recall

to solve $f(x)=0$, Start from x_0

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\hat{f}(x)} + \cdots$$

$\hat{f}(x)$ -- ← linear polynomial (polynomial of degree 1)

we instead solve

$\hat{f}(x)=0 \Rightarrow$ This gives x_1

Newton's method

Newton: approximate $f(x)$ by a linear polynomial

↓

generalize this to systems of equations.

$n=2$

$$\begin{cases} f(x,y)=0 \\ g(x,y)=0. \end{cases}$$

start with an approximation (x_0, y_0)

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)$$

$$+ \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$$

$\hat{f}(x,y)$

$$+ \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x_0,y_0)(x-x_0)^2$$

$$+ \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)(x-x_0)(y-y_0)$$

$$+ \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(x_0,y_0)(\cancel{x}\, y-y_0)^2$$

$$+ \cdots$$

$$g(x,y) = \underbrace{g(x_0,y_0) + \frac{\partial g}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial g}{\partial y}(x_0,y_0)(y-y_0)}_{\hat{g}(x,y)} + \cdots\cdots$$

Newton's method is to

Solve $\begin{cases} \hat{f}(x,y)=0 \\ \hat{g}(x,y)=0 \end{cases}$ to get (x_1, y_1)

To better see the structure, we can work with the vector form.

let $\underline{F}(x,y) = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$

$$\underbrace{D\underline{F}(x,y)}_{\text{write it as } J_F(x,y)} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix}$$ Jacobian.

incremental part

Approximation step:

$$\underline{F}(x,y) \approx \underbrace{\underline{F}(x_0,y_0) + J_F(x_0,y_0)\overbrace{\begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}}^{s}}_{=0\ \text{(Newton's method)}}$$

Newton's method: Start with an initial $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, for $k = 0, 1, 2, \ldots$

Solve $\underline{s} \in \mathbb{R}^2$ from

$$\begin{cases} J_F(x_k, y_k)\, \underline{s} = -\underline{F}(x_k, y_k) \\ \begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} + \underline{s} \quad \leftarrow \text{incremental} \end{cases}$$

Stop if $\|\underline{s}\| < \mathrm{Tol}$, or $k > K_{max}$

Core: linearization ←

Back to the example

$$\begin{cases} f(x, y) = y - x^3 \\ g(x, y) = x^2 + y^2 - 1 \end{cases}$$

$$J_F(x, y) = \begin{pmatrix} -3x^2 & 1 \\ 2x & 2y \end{pmatrix}$$

initial $x_0 = 1$, $y_0 = 1$

$$J_F(x_0 = 1, y_0 = 1) = \begin{pmatrix} -3 & 1 \\ 2 & 2 \end{pmatrix}$$

$$\underline{F}(x_0, y_0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

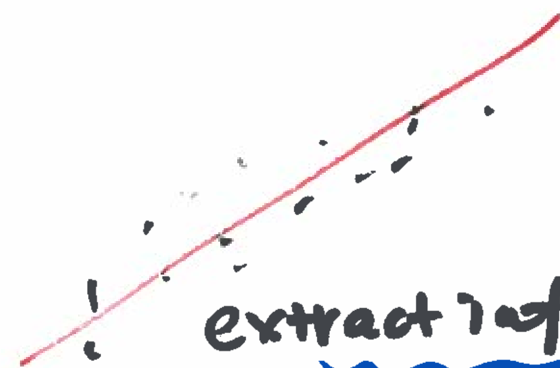
Step 1 of Newton's method.

$$\text{Solve } \underbrace{\begin{pmatrix} -3 & 1 \\ 2 & 2 \end{pmatrix}}_{= J_F(x_0, y_0)} \underline{s} = -\underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{= \underline{F}(x_0, y_0)}$$

$$\underline{s} = \begin{pmatrix} -\frac{1}{8} \\ -\frac{3}{8} \end{pmatrix}$$

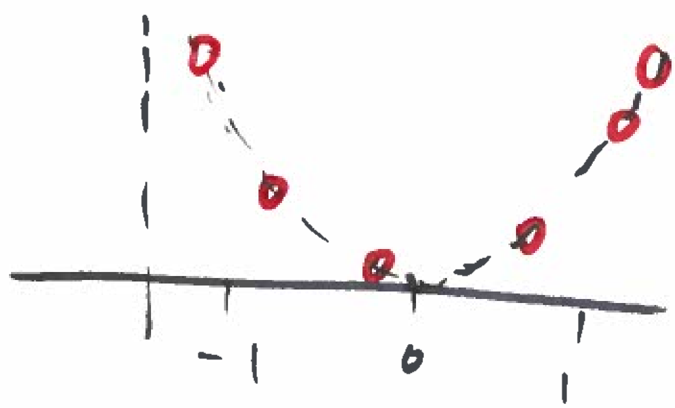
$$\Rightarrow \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \underline{s} = \begin{pmatrix} 7/8 \\ 5/8 \end{pmatrix}$$

§ 3 Interpolation and data fitting



extract information of given data

$$f(x) = \frac{1}{3} + \sum_{n=1}^{\infty} \frac{4(-1)^n}{n^{4/3}\pi^2} \cos(n\pi x)$$

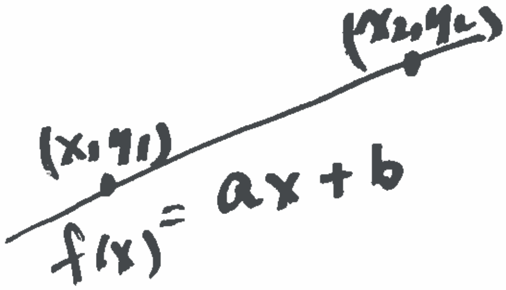


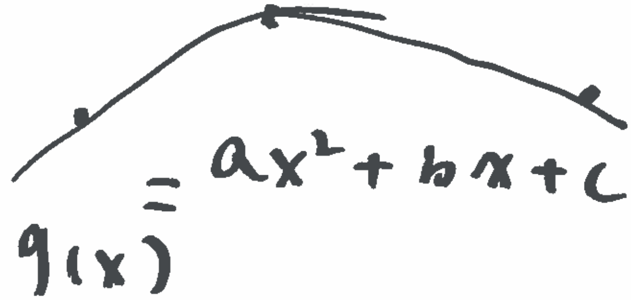
$f(x) \approx \hat{f}(x)$ based 6 points on the graph

$\hat{f}(x)$: simpler form, an approximation of $f(x)$.

- extract information of data
- approximate functions

§3-1 Interpolation : global





In general, given $(x_1, y_1), (x_2, y_2), \dots, (x_{n+1}, y_{n+1})$

Def: A function $y = p(x)$ interpolates the points given above if
$$y_j = p(x_j) \quad j = 1, 2, \dots, n+1.$$

For now we assume $p(x)$ is a polynomial of degree n

(we know uniqueness and existence of such $p(x)$ when $n = 2$, and $n=3$)

Problem:

$$\begin{cases} \text{Given} & (x_j, y_j) = \quad j = 1, 2 \dots n+1 \\ \text{find} & P_n(x), \text{ such that } P_n(x_j) = y_j \quad j = 1 \dots n+1 \end{cases}$$

$\uparrow$ polynomial of degree n

Approach: direct method

look for

$$P_n(x) = a_0 + a_1 x + \cdots + a_n x^n$$

such that $P_n(x_j) = y_j \quad j = 1, \dots n+1$

$$\Leftrightarrow \begin{cases} a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_n x_1^n = y_1 \\ a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_n x_2^n = y_2 \\ \vdots \\ a_0 + a_1 x_{n+1} + a_2 x_{n+1}^2 + \cdots + a_n x_{n+1}^n = y_{n+1} \end{cases}$$

In matrix-vector form:

$$\underline{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \quad \underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n+1} \end{bmatrix}, \quad \text{we have}$$

$$\boxed{A \underline{a} = \underline{y}} \qquad A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdot & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ & \vdots & & & \\ 1 & x_{n+1} & x_{n+1}^2 & \cdots & x_{n+1}^n \end{pmatrix}$$

$\checkmark \quad ? \quad \checkmark$

$\checkmark$

Vandermonde matrix

$A\underline{a} = \underline{y}$, A is square

unique solvability $\Leftrightarrow$ A is invertible

$\Leftrightarrow$ $\det(A) \neq 0$.

Lemma: $\det(A) = \prod_{1 \le i < j \le n+1} (x_j - x_i)$

when $n=1$

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} \qquad \det(A) = x_2 - x_1$$

when $n=2$

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{pmatrix} \qquad \det(A) = (x_3-x_1)(x_2-x_1)(x_2-x_1)$$

Theorem: Given $\{(x_j, y_j)\}_{j=1}^{n+1}$ the interpolating polynomial $p_n(x)$ exists uniquely $\Leftrightarrow$ $\{x_j\}_{j=1}^{n+1}$ are distinct.