Plan

1. Orthogonal diagonalization (continued from Lecture 5)
2. Unconstrained optimization
3. Convex and concave functions

Recall:

A $n\times n$ i) An orthogonal diagonalization of A is a diagonalization $P^{-1}AP = D$ such that P is an orthogonal matrix, i.e. $P^{-1} = P^T$.

$P = (\underline{v_1}|\underline{v_2}|\ldots|\underline{v_n})$: $P^{-1} = P^T \iff \begin{cases} (a)\ \underline{v_i}\cdot\underline{v_j} = 0 \text{ for } i \neq j \\ (b)\ \|v_i\| = 1 \end{cases}$

ii) There is an orth. diag. of A $\iff$ A is symmetric

① Orthogonal diagonalization (cont'd)

Quadratic form in n variables: $f(\underline{x}) = \underline{x}^T A \underline{x}$ $\quad \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ $\quad A$ $n\times n$ symm.

Fact: A has orth. diagonalization: $P^T A P = D$

Change of variables: $\underline{x} = P\underline{u}$ $\quad |P^{-1}\cdot$ $\quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = P\cdot\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ or $\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = P^T\cdot\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

$P^T\cdot\underline{x} = \underline{u}$

$$f(\underline{x}) = \underline{x}^T A \underline{x} = (P\underline{u})^T A (P\underline{u}) = \underline{u}^T \underbrace{P^T A P}_{D} \underline{u} = \underline{u}^T D \underline{u}$$
$$= \lambda_1 u_1^2 + \lambda_2 u_2^2 + \ldots + \lambda_n u_n^2$$

Ex: $f(x,y,z) = 4x^2 - 2xy - 6xz + 2y^2 - 2yz + 4z^2$

$\underline{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ $\quad A = \begin{pmatrix} 4 & -1 & -3 \\ -1 & 2 & -1 \\ -3 & -1 & 4 \end{pmatrix}$

Eigenvalues: $-\lambda^3 + c_1\lambda^2 - c_2\lambda + c_3 = 0$

$-\lambda^3 + 10\lambda^2 - 21\lambda + 0 = 0$

$\begin{cases} c_1 = tr(A) = 4 + 2 + 4 = 10 \\ c_2 = \quad = 7 + 7 + 7 = 21 \\ c_3 = |A| = 0 \end{cases}$

$-\lambda(\lambda^2 - 10\lambda + 21) = 0$

$\lambda_1 = 0 \quad \lambda_2 = 3 \quad \lambda_3 = 7$

$R1 + R2 + R3 = \underline{0}$

Viète's formula:

$\lambda^2 - a\lambda + b = 0$

$\lambda = \lambda_1$ and λ_2 are the soln's

$\Updownarrow$

$\lambda_1 + \lambda_2 = a \,,\quad \lambda_1 \cdot \lambda_2 = b$

Ex:

$\lambda^2 - 10\lambda + 21 = 0$

$\lambda_1 + \lambda_2 = 10$

$\lambda_1 \cdot \lambda_2 = 21$

Ex: (cont'd)

$$f(x,y,z) = 4x^2 - 2xy - 6xz + 2y^2 - 2yz + 4z^2 = \underline{0 \cdot u^2 + 3 \cdot v^2 + 7 w^2}$$

We immediately see that $\underline{f_{min} = 0}$

How to compute P

(a) $\underline{v_i} \cdot \underline{v_j} = 0 \ (i \neq j) \rightarrow$ Always ok if $\underline{v_i}$ and $\underline{v_j}$ corr. to different eigenvalues

(b) $\|\underline{v_i}\| = 1$.

E_0: $\begin{pmatrix} 4 & -1 & -3 \\ -1 & 2 & -1 \\ -3 & -1 & 4 \end{pmatrix} \rightsquigarrow \underline{v_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

E_3: $\begin{pmatrix} 1 & -1 & -3 \\ -1 & -1 & -1 \\ -3 & -1 & 1 \end{pmatrix} \quad \underline{v_2} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

E_7: $\begin{pmatrix} -3 & -1 & -3 \\ -1 & -5 & -1 \\ -3 & -1 & -3 \end{pmatrix} \quad \underline{v_3} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

$\rightarrow \underline{v_1}' = \frac{1}{\|\underline{v_1}\|} \underline{v_1} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$\underline{v_2}' = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

$\underline{v_3}' = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

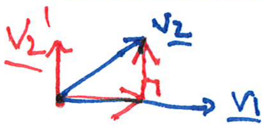
$$P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{pmatrix}$$

Change of variables:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \underbrace{\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}}_{= P^T} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\boxed{\begin{aligned} u &= \tfrac{1}{\sqrt{3}}(x + y + z) \\ v &= \tfrac{1}{\sqrt{6}}(x - 2y + z) \\ w &= \tfrac{1}{\sqrt{2}}(-x \quad + z) \end{aligned}}$$

If $\underline{v_1}, \underline{v_2}$ is in E_λ and $\underline{v_1} \cdot \underline{v_2} \neq 0$:



$$\underline{v_2}' = \underline{v_2} - \mathrm{proj}_{\underline{v_1}}(\underline{v_2}) = \underline{v_2} - \frac{\underline{v_2} \cdot \underline{v_1}}{\underline{v_1} \cdot \underline{v_1}} \cdot \underline{v_1}$$

(2) Unconstrained optimization

max/min $f(x_1, x_2, \ldots, x_n) = f(\underline{x})$

Assume that f is "nice" (C^2, i.e all partial derivatives and second order partial derivatives exists and are cont.)

Ex: $f(x,y,z) = 4x^2 - 2xy - 6xz + 2y^2 - 2yz + 4z^2$
$= 0 \cdot u^2 + 3 \cdot v^2 + 7w^2$

$f'_x = 8x - 2y - 6z$
$f'_y = -2x + 4y - 2z$
$f'_z = -6x - 2y + 8z$

first order partial derivatives

$$H(f) = \begin{pmatrix} f''_{xx} & f''_{xy} & f''_{xz} \\ f''_{yx} & f''_{yy} & f''_{yz} \\ f''_{zx} & f''_{zy} & f''_{zz} \end{pmatrix} = \begin{pmatrix} 8 & -2 & -6 \\ -2 & 4 & -2 \\ -6 & -2 & 8 \end{pmatrix}$$

Hessian matrix

Fact: $H(f)$ is symmetric (if f is "nice")

Method:

(1) Find all stationary pts of f: $\boxed{f'_x = f'_y = f'_z = 0}$ (FOC)
"Candidate pts"

Fact: If $\underline{x}^* = (x^*, y^*, z^*)$ is max/min, then it is stationary

Ex: $f = x^3 + y^3 + z^3 - 3x - 3y - 3z + 5$

$f'_x = 3x^2 - 3 = 0 \qquad x^2 = 1 \qquad x = \pm 1$
$f'_y = 3y^2 - 3 = 0 \qquad y^2 = 1 \qquad y = \pm 1$
$f'_z = 3z^2 - 3 = 0 \qquad z^2 = 1 \qquad z = \pm 1$

Stationary pts:

$(1,1,1)$ $f = -1$ min?
$(1,1,-1), (1,-1,1), (-1,1,1)$ $f = 3$
$(1,-1,-1), (-1,1,-1), (-1,-1,1)$ $f = 7$
$(-1,-1,-1)$ $f = 11$ max?

Definitions:

$f(\underline{x}) = f(x_1, \ldots, x_n)$ function in n variables

$\underline{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)$ a particular point

Defn.

$\underline{x}^*$ is a maximum pt (global maximum point) for f if $f(\underline{x}^*) \geq f(\underline{x})$ for all points $\underline{x}$, and $f_{max} = f(\underline{x}^*)$ is then the maximum value of f

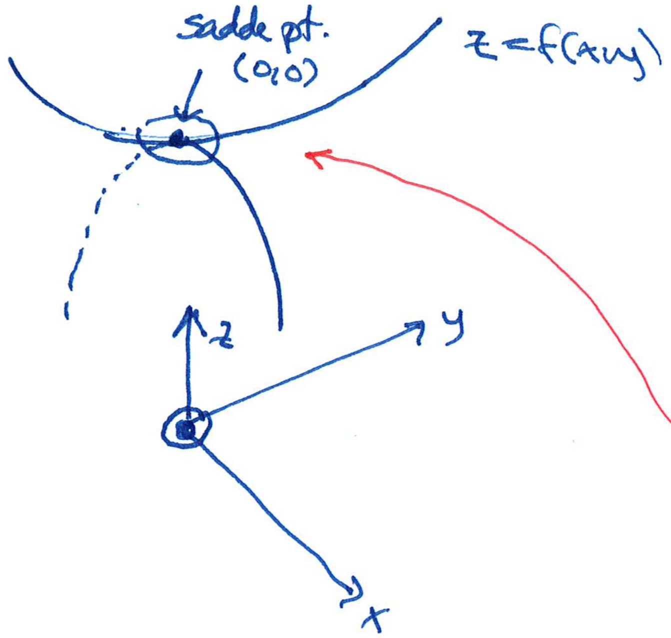
$\underline{x}^*$ is a local maximum pt. for f if $f(\underline{x}^*) \geq f(\underline{x})$ for all points $\underline{x}$ close to $\underline{x}^*$

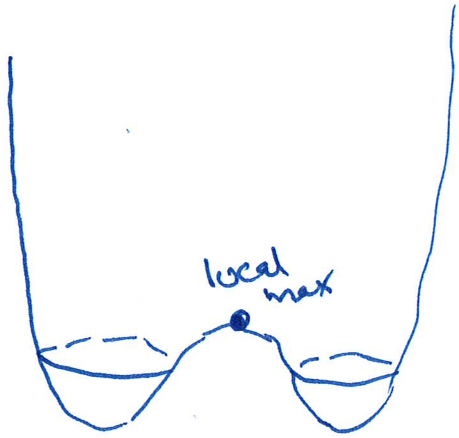
$\underline{x}^*$ is a minimum pt (global minimum pt) for f if $f(\underline{x}^*) \leq f(\underline{x})$ for all points $\underline{x}$, and $f_{min} = f(\underline{x}^*)$ is then the minimum value of f

$\underline{x}^*$ is a local minimum pt for f if $f(\underline{x}^*) \leq f(\underline{x})$ for all pts. $\underline{x}$ close to $\underline{x}^*$.

$\underline{x}^*$ is a saddle pt for f if $\underline{x}^*$ is a stationary pt. of f that is neither a local max nor a local min.

The range V_f of f is the set of all function values $f(\underline{x})$ of f, that is; $V_f = \{ f(\underline{x}) : \underline{x}$ in D_f = domain of defn. of $f \}$.





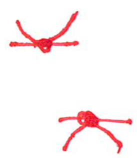
(2) <u>Classify the stationary points</u> as local max, local min, or saddle pts using the second derivative test:

Result: Let $\underline{x}^*$ be a stationary pt of f. Then we have:

$H(f)(\underline{x}^*)$ pos. defn. $\Rightarrow$ $\underline{x}^*$ local min

$H(f)(\underline{x}^*)$ neg defn. $\Rightarrow$ $\underline{x}^*$ local max

— " — indefn. $\Rightarrow$ $\underline{x}^*$ saddle pt.



<u>Ex:</u> (cont'd) $f = x^3 + y^3 + z^3 - 3x - 3y - 3z + 5$

$f'_x = 3x^2 - 3$

$f'_y = 3y^2 - 3$

$f'_z = 3z^2 - 3$

$$H(f) = \begin{pmatrix} 6x & 0 & 0 \\ 0 & 6y & 0 \\ 0 & 0 & 6z \end{pmatrix}$$

$(1,1,1)$: $H(f)(1,1,1) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ pos. defn. $\Downarrow$ $(1,1,1)$ local min

$(-1,-1,-1)$: $= \begin{pmatrix} -6 & & \\ & -6 & \\ & & -6 \end{pmatrix}$ neg. defn. $\Downarrow$ $(-1,-1,-1)$ local max

all other stat. pts $\}$ <u>saddle pts</u>

(3) Determine if best candidate for max/min is actually max/min.

Ex: $f = x^3 + y^3 + z^3 - 3x - 3y - 3z + 5$

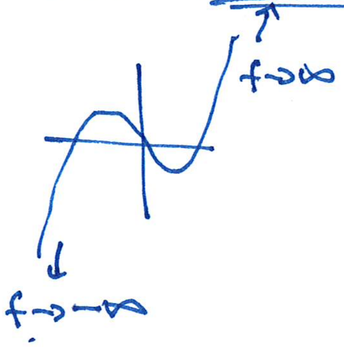
$f(1,1,1) = -1$ local min → Best cand. for min. Not min

$f(-1,-1,-1) = 11$ local max → Best cand. for max. Not max

Min: Try $f(x,0,0) = x^3 - 3x + 5$

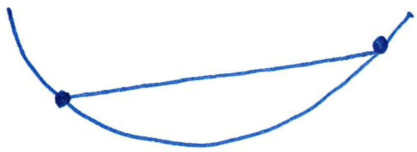
$f(-10,0,0) = -1000 + 30 + 5 = -965$

$f(10,0,0) = 1000 - 30 + 5 = 975$

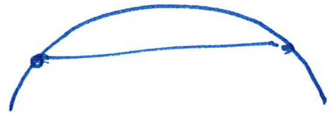


Concl: No (global) max/min.

③ Convex / concave functions



Convex



concave

Criterion for convex / concave function:

If $H(f)(\underline{x})$ is positive semidefin. for all pts $\underline{x}$, then f is convex
If —‖— negative —‖— —‖— f is concave

In fact:

$$\boxed{\begin{array}{l} H(f)(\underline{x}) \text{ pos. semidefin. for all } \underline{x} \iff f \text{ convex} \\ H(f)(\underline{x}) \text{ neg. semidefin. —‖—} \iff f \text{ concave} \end{array}}$$

Ex: $f = x^3 + y^3 + z^3 - 3x - 3y - 3z + 5$

$$H(f) = \begin{pmatrix} 6x & 0 & 0 \\ 0 & 6y & 0 \\ 0 & 0 & 6z \end{pmatrix}$$

pos. defin. at $(1,1,1)$
neg. defin. at $(-1,-1,-1)$

not convex, not concave

$$f(x,y,z) = \left\{ \begin{array}{l} 4x^2 - 2xy - 6xz \\ + 2y^2 - 2yz + 4z^2 \end{array} \right\} = \underline{x}^T A \underline{x}$$

$$H(f) = 2A = \begin{pmatrix} 8 & -2 & -6 \\ -2 & 4 & -2 \\ -6 & -2 & 8 \end{pmatrix}$$

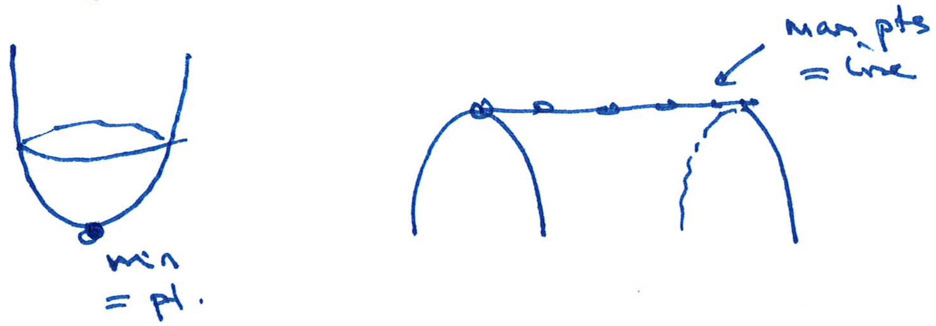
A pos. semidefin. $\lambda = 0, 3, 7$
$H(f) = 2A$ pos. semidefin. $\lambda = 0, 6, 14$

f **convex**

Result: convex optimization

i) If f is convex and $\underline{x}^*$ stationary pt of f, then it is min

ii) — " — concave — " — then it is max



Ex:
$$f(x,y,z) = \left.\begin{array}{l} 4x^2 - 2xy - 6xz \\ + 2y^2 - 2yz + 4z^2 \end{array}\right\} = \underline{x}^T A \underline{x} + B\underline{x} + C$$
$$+ 14x - 14z + 5$$

$B = (14 \quad 0 \quad -14)$

$B\underline{x} = (14 \quad 0 \quad -14) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

In this case:

$$H(f) = 2A$$

$$f'(\underline{x}) = \begin{pmatrix} f'_{x_1} \\ f'_{x_2} \\ \vdots \\ f'_{x_n} \end{pmatrix} = 2A\underline{x} + B^T$$

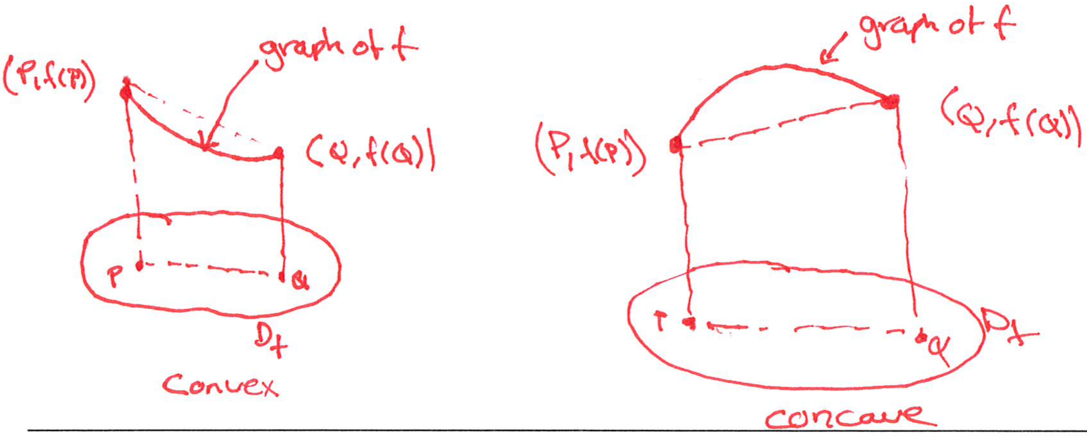
More on defn. of convex/concave fn's:

Let $f(x_1, x_2, \dots, x_n)$ be a fn. defined on a subset D_f of $\mathbb{R}^n$. Then D_f is called a <u>convex set</u> if the line segment $[P,Q]$ from a point $P=(p_1,\dots,p_n)$ in D_f to a pt $Q=(q_1,\dots,q_n)$ in D_f is contained in D_f for all P, Q in D_f.

<u>Ex:</u>

To define whether f is convex/concave, D_f must be a convex set.

<u>Defn:</u> Let $f(x)$ be defined on a convex set D_f.
We say that f is <u>convex</u> (concave) if for any pts P, Q in D_f, the line segment from $(P, f(P))$ to $(Q, f(Q))$ lies on or above (on or below) the graph of f.



Ex: $f(x,y,z) = 4x^2 - 2xy - 6xz + 2y^2 - 2yz + 4z^2 + 14x - 14z + 3$

$= \underline{x}^T A \underline{x} + B\underline{x} + C$

$$A = \begin{pmatrix} 4 & -1 & -3 \\ -1 & 2 & -1 \\ -3 & -1 & 4 \end{pmatrix}, \quad B = (14\ 0\ -14), \quad C = (3)$$

Stat. pts: $f'(\underline{x}) = \begin{pmatrix} f'_x \\ f'_y \\ f'_z \end{pmatrix} = 2A\underline{x} + B^T = \underline{0}$

$2A\underline{x} = -B^T$

$A\underline{x} = -\frac{1}{2}B^T$

$$\begin{array}{ccc|c} A & & & -\tfrac{1}{2}B^T \\ \downarrow & & & \downarrow \end{array}$$

$$\left(\begin{array}{ccc|c} 4 & -1 & -3 & -7 \\ -1 & 2 & -1 & 0 \\ -3 & -1 & 4 & 7 \end{array}\right) \to \left(\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 4 & -1 & -3 & -7 \\ -3 & -1 & 4 & 7 \end{array}\right) \begin{matrix} \\ 4 \\ -3 \end{matrix} \to \left(\begin{array}{ccc|c} -1 & 2 & -1 & 0 \\ 0 & 7 & -7 & -7 \\ \cancel{0} & \cancel{-7} & \cancel{7} & \cancel{7} \end{array}\right)$$

z free

$-x + 2y - z = 0$

$7y - 7z = -7$

$\Rightarrow 7y = 7z - 7 \Rightarrow y = \underline{z-1}$

$-x + 2(z-1) - z = 0 \Rightarrow x = \underline{z-2}$

Stat. pts: $(x,y,z) = \underline{(z-2, z-1, z)}$, ($z$ free)

Convex/Concave: $H(f) = 2A$, $A = \begin{pmatrix} 4 & -1 & -3 \\ -1 & 2 & -1 \\ -3 & -1 & 4 \end{pmatrix}$ $\left.\begin{matrix} D_1 = 4 \\ D_2 = 7 \\ D_3 = 0 \end{matrix}\right\} \Rightarrow$ A pos. semidef. by RKEC (rk A = 2)

f convex

$\Rightarrow (z-2, z-1, z)$ are min. pts of f

$f(z-2, z-1, z) = f(-2,-1,0) = 16 - 4 + 2 - 28 + 3 = -11$
(z=0)

$f_{min} = \underline{\underline{-11}}$ min. value

at

$(x,y,z) = \underline{(z-2, z-1, z)}$ for all z — min. pts