Plan

1 Orthogonal diagonalization
2 Unconstrained optimization
3 Convex and concave functions

① Orthogonal diagonalization :

Review + material that I didn't have time for last week

A $n \times n$ matrix

Defn: An orthogonal diagonalization of A is a diagonalization $P^{-1}AP = D$ where $P^{-1} = P^T$.

$$D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \qquad P = \left(\underline{v_1} \,|\, \underline{v_2} \,|\, \cdots \,|\, \underline{v_n} \right)$$

diagonal matrix (eigenvalues) (eigenvectors)

$\|v_1\| = 1, \|v_2\| = 1, \ldots$

Note: $P^{-1} = P^T \iff$ i) $\underline{v_1} \cdot \underline{v_1} = \underline{v_2} \cdot \underline{v_2} = \cdots = \underline{v_n} \cdot \underline{v_n} = 1$ and ii) $\underline{v_i} \cdot \underline{v_j} = 0$ for all i, j with $i \neq j$

$\underline{v_i} \perp \underline{v_j}$

When? A orthogonal diagonalizable $\iff$ A is symmetric

How? a) If $\underline{v}$ is an eigenvector, then $\frac{1}{\|\underline{v}\|} \cdot \underline{v}$ satisfies (i).

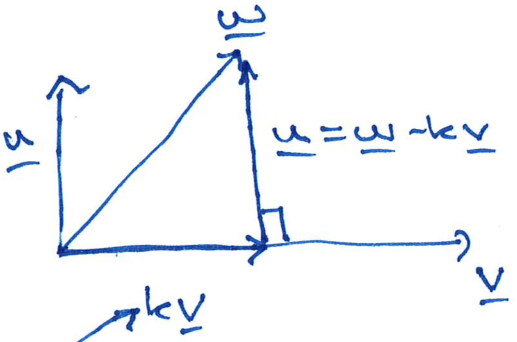
$\|\underline{v}\|$ = length of $\underline{v}$ = $\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$

b) If A is symmetric and $\underline{v}, \underline{w}$ are eigenvectors with different eigenvalues, then $\underline{v} \cdot \underline{w} = 0$, and (ii) is satisfied (for $\underline{v}, \underline{w}$)

$\underline{v} \perp \underline{w} \iff \underline{v} \cdot \underline{w} = 0 \iff v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = 0$

c) If $\underline{v}, \underline{w}$ are eigenvectors with the same eigenvalue λ:

We don't necessarily have $\underline{v} \cdot \underline{w} = 0$



$$\underline{w} = k\underline{v} + \underline{u}$$

$$\Rightarrow \underline{u} = \underline{w} - k\underline{v}$$

$$\underline{v} \cdot \underline{u} = 0 \quad \leftarrow \quad \underline{v} \perp \underline{u}$$

$$\mathrm{proj}_{\underline{v}}(\underline{w}) = k\underline{v}$$

$$= \left(\frac{\underline{v}\cdot\underline{w}}{\underline{v}\cdot\underline{v}}\right)\underline{v}$$

$$\underline{v} \cdot (\underline{w} - k\underline{v}) = 0$$

$$\underline{v} \cdot \underline{w} - k(\underline{v}\cdot\underline{v}) = 0$$

$$k = \frac{\underline{v}\cdot\underline{w}}{\underline{v}\cdot\underline{v}}$$

$\Rightarrow$ We replace $\underline{v}, \underline{w}$ with $\underline{v}, \underline{u}$

Why: Let $q(\underline{x}) = \underline{x}^T A \underline{x}$ be a quadratic form in n var's. with $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and A ~~is~~ an $n \times n$ symmetric matrix.

$\Rightarrow$ A has an orthogonal diagonalization: $\quad P^T A P = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & \ddots & \lambda_n \end{pmatrix}$

Linear change of variables: $\quad \underline{x} = P \cdot \underline{u} \iff \underline{u} = P^T \underline{x}$

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = P^T \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$q(\underline{x}) = \underline{x}^T A \underline{x} = (P\underline{u})^T A (P\underline{u})$$

$$= \underline{u}^T \underbrace{P^T A P}\underline{u} = \underline{u}^T D \underline{u}$$

$$= \lambda_1 \cdot u_1^2 + \lambda_2 u_2^2 + \dots + \lambda_n u_n^2$$

Ex: $q(x,y,z) = \underline{2x^2} + \underline{2xy} + \underline{2xz} + \underline{2y^2} + \underline{2yz} + \underline{2z^2}$

$$= (x\ y\ z)\cdot \underbrace{\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underline{x}^T A \underline{x}$$

Find orthogonal diagonalization of A:

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = 0 :$$

$\underline{\lambda = 1}$ $\begin{pmatrix} 1 & 1 & 1 \\ \cancel{1} & \cancel{1} & \cancel{1} \\ \cancel{1} & \cancel{1} & \cancel{1} \end{pmatrix}$ $\Rightarrow$ $\lambda = 1$ is an eigenvalue and $\dim E_1 = 2$

$1 \leq \underset{2}{\overset{\parallel}{\dim E_1}} \leq m$

$\Rightarrow m \geq 2$

$\underline{\lambda_1 = \lambda_2 = 1}$

$\lambda_1 + \lambda_2 + \lambda_3 = tr(A) = 6$

$1 + 1 + \lambda_3 = 6$ $\quad \underline{\lambda_3 = 4}$

$\Rightarrow q(\underline{x}) = 1\cdot u^2 + 1\cdot v^2 + 4w^2$
$= \underline{u^2 + v^2 + 4w^2}$

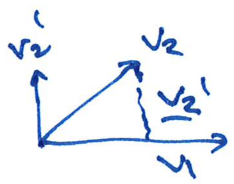
Eigenvectors:

$\underline{\lambda = 1}$: $x + y + z = 0$
y, z free

$x = -y - z$
$y = y$
$z = z$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = y\cdot \underset{\underline{v_1}}{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}} + z\cdot \underset{\underline{v_2}}{\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}}$$

$\underline{v_1}\cdot\underline{v_2} = 1 \neq 0$
$\underline{v_1}\cdot\underline{v_1} = 2$



$$\text{proj}_{\underline{v_1}}(\underline{v_2}) = \left(\frac{\underline{v_1}\cdot\underline{v_2}}{\underline{v_1}\cdot\underline{v_1}}\right)\cdot \underline{v_1} = \frac{1}{2}\underline{v_1}$$

$$\underline{v_2}' = \underline{v_2} - \frac{1}{2}\underline{v_1} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2}\cdot\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$$

New vectors in E_1: $\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$

$\underline{\lambda = 4}$: $\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \to \dots \to$ $\quad \underline{v_3} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $\quad$ New vector in E_4: $\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$$P = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} ; \quad \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$q(x,y,z) = 2x^2 + 2xy + 2xz + 2y^2 + 2yz + 2z^2 = \underline{u^2 + v^2 + 4w^2}$$

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

(2) Unconstrained optimization in n variables

$$\boxed{\max / \min \quad f(x_1, x_2, x_3, \ldots, x_n) = f(\underline{x})}$$

Ex: $f(x,y,z) = x^3 + y^3 + z^3 - 3xyz$, $D = D_f = \mathbb{R}^3$

$f(x,y,z) = 2x^2 + 2xy + 2xz + 2y^2 + 2yz + 2z^2$, $D = D_f = \mathbb{R}^3$

$f(x,y,z) = \ln(x - y + z)$, $D: x - y + z > 0$

Note: $f(x,y,z) = u^2 + v^2 + 4w^2$ with $\begin{cases} u = \frac{1}{\sqrt{2}}(-x+y) \\ v = \frac{1}{\sqrt{6}}(-x-y+2z) \\ w = \frac{1}{\sqrt{3}}(x+y+z) \end{cases}$

$\Rightarrow f_{min} = \underline{\underline{0}}$ (no max)

$$\boxed{\text{Assumption: } \underbrace{f \text{ is } C^2}_{\text{first and second order partial der. of } f \text{ are "nice"}} \text{ on } \underbrace{\text{an open subset } D \text{ of } \mathbb{R}^n}_{\text{"}D = \mathbb{R}^n \text{ or } D \text{ is given by } <, >, \neq\text{"}}}$$

Review of some definitions: $f(\underline{x})$ defined on $D = D_f$

$\underline{x}^*$ is a max (or global max.) of f if $f(\underline{x}^*) \geq f(\underline{x})$ for all $\underline{x}$ in D

— " — min (or global min.) — " — $f(\underline{x}^*) \leq f(\underline{x})$ — " —

If $\underline{x}^*$ is a max/min of f, we write: $f_{max} = f(\underline{x}^*)$ or $f_{min} = f(\underline{x}^*)$

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

— " — local max — " — $f(\underline{x}^*) \geq f(\underline{x})$ — " —

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

Range: The range of f is the set of possible function values of f.

If f has global max f_{max} and global min f_{min}, then the range of f is:

$$V_f = \{ f(\underline{x}) : \underline{x} \text{ in } D_f \} = [f_{min}, f_{max}]$$

Method: max/min $f(\underline{x})$

① Stationary points: $\boxed{f'_{x_1} = f'_{x_2} = \dots = f'_{x_n} = 0}$ (FOC)

Ex: $f(x,y,z) = x^3 + y^3 + z^3 - 3(x+y+z)$

FOC
$$\begin{cases} f'_x = 3x^2 - 3 = 0 & \frac{3x^2}{3} = \frac{3}{3} \quad x^2 = 1 \quad \underline{x = \pm 1} \\ f'_y = 3y^2 - 3 = 0 & y^2 = 1 \quad \underline{y = \pm 1} \\ f'_z = 3z^2 - 3 = 0 & z^2 = 1 \quad \underline{z = \pm 1} \end{cases}$$

Stationary pts: $(x,y,z) = (1,1,1)$ $f = -6$

min? → $(1,1,1)$

$(-1,1,1)$, $(1,-1,1)$, $(1,1,-1)$ $f = -2$

$(-1,-1,1)$, $(-1,1,-1)$, $(1,-1,-1)$, $f = 2$

max? → $(-1,-1,-1)$ $f = 6$

Fact: If $\underline{x}^*$ is a max/min, it must be a stationary pt.

② Local classification: $H(f) = \begin{pmatrix} f''_{xx} & f''_{xy} & f''_{xz} \\ f''_{xy} & f''_{yy} & f''_{yz} \\ f''_{xz} & f''_{yz} & f''_{zz} \end{pmatrix}$

Hessian of f

Note: $H(f)$ is a symmetric $n \times n$ matrix.

Second derivative test:

If $\underline{x}^*$ is a stationary pt of f, consider $H(f)(\underline{x}^*)$:

$H(f)(\underline{x}^*)$ pos. defn. ⇒ $\underline{x}^*$ is a local ~~max~~ min.

— " — neg. defn. ⇒ $\underline{x}^*$ is a local max.

— " — indefn. ⇒ $\underline{x}^*$ is a saddle pt.

Ex: $f'_x = \underline{3x^2-3}$ $\quad f''_{xx}=6x \quad f''_{xy}=0 \quad f''_{xz}=0$

$f'_y = \underline{3y^2-3}$ $\quad f''_{xy}=0 \quad f''_{yy}=6y \quad f''_{yz}=0$

$f'_z = \underline{3z^2-3}$ $\quad f''_{xz}=0 \quad f''_{yz}=0 \quad f''_{zz}=6z$

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

$\underline{(1,1,1)}$: $H(f)(1,1,1)=\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ $\lambda_1=6$, $\lambda_2=6$, $\lambda_3=6$ $\quad$ $\underline{\text{pos. defn}}$ $\Rightarrow (1,1,1)$ $\underline{\text{local min}}$

$\underline{(-1,1,1)}$: $H(f)(-1,1,1)=\begin{pmatrix} -6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ $\lambda_1=-6$, $\lambda_2=6$, $\lambda_3=6$ $\quad$ $\underline{\text{indefn.}}$ $\Rightarrow (-1,1,1)$ saddle pt

$\underline{(-1,-1,-1)}$: $H(f)(-1,-1,-1)=\begin{pmatrix} -6 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & -6 \end{pmatrix}$

$\lambda_1=-6$, $\lambda_2=-6$, $\lambda_3=-6$ $\quad$ neg. defn. $\Rightarrow (-1,-1,-1)$ local max

$\underline{\text{Candidate for min:}}$ Best candidate $f(1,1,1)=-6$ local min

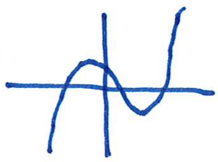
$\underline{\text{Candidate for max:}}$ — " — $f(-1,-1,-1)=6$ local max

$\Rightarrow$ $\boxed{\text{No global max/min}}$

③ Determine if candidates for max/min $\underline{\text{are max/min}}$:

$f(x,y,z)=x^3+y^3+z^3-3(x+y+z)$

$f(x,0,0)=x^3-3x$



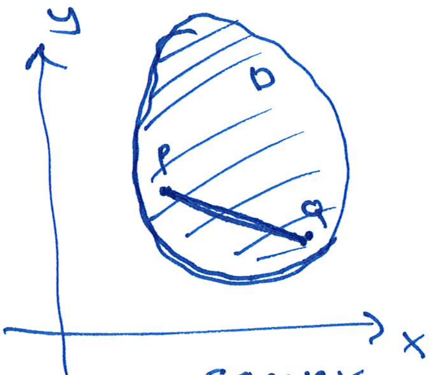
$f(3,0,0)=27-9=18>6$

$f(-3,0,0)=-27+9=-18<-6$

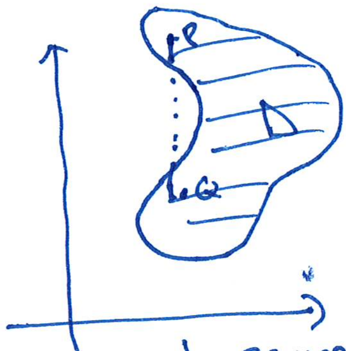
③ Convex and concave functions

Assumption: f is C^2 on an open convex set D

Note: A subset D of $\mathbb{R}^n$ is convex if the following condition is satisfied: For any two points P, Q in D, the line segment $[P,Q]$ from P to Q is in D.



convex



not convex

Defn:

f is convex if $H(f)(\underline{x})$ is positive semidefn. for all $\underline{x}$

f is concave if $H(f)(\underline{x})$ is negative semidefn. — " —

Ex: $f(x,y,z) = x^3+y^3+z^3 - 3(x+y+z)$

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

$\lambda_1 = 6x \geq 0$ $\quad \leq 0$

$\lambda_2 = 6y \geq 0$ $\quad \leq 0$

$\lambda_3 = 6z \geq 0$ $\quad \leq 0$

f is not convex, not concave $\Leftarrow$ No. No.

Result:

If f is convex, then any stationary pt of f is a global min.

If f is concave, then any stationary pt of f is a global max

Ex: max/min $f(x,y,z,w) = \overbrace{x^2+y^2+z^2+w^2+xy+yz+zw}^{\text{quadratic form}}$

$$\underbrace{-x-y+z+w}_{\text{linear form}}\ \underbrace{-1}_{\text{const.}}$$

Note: $f(\underline{x}) = \underline{x}^T A \underline{x} + B\underline{x} + C$

$$A = \begin{pmatrix} 1 & 1/2 & 0 & 0 \\ 1/2 & 1 & 1/2 & 0 \\ 0 & 1/2 & 1 & 1/2 \\ 0 & 0 & 1/2 & 1 \end{pmatrix} \quad \underline{x} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

$f(x) = ax^2+bx+c$

$f'(x) = 2ax+b$

$f''(x) = 2a$

$B = \begin{pmatrix} -1 & -1 & 1 & 1 \end{pmatrix}$

$C = -1$

Fact: $f(\underline{x}) = \underline{x}^T A \underline{x} + B\underline{x} + C \leftarrow$ polynomial fn. of degree 2.

$\Downarrow$

$$\boxed{\begin{array}{l} 1)\ f'(\underline{x}) = \begin{pmatrix} f'_{x_1} \\ f'_{x_2} \\ \vdots \\ f'_{x_n} \end{pmatrix} = 2A\underline{x} + B^T \\ 2)\ H(f) = 2A \end{array}}$$

$$A = \begin{pmatrix} 1 & 1/2 & 0 & 0 \\ 1/2 & 1 & 1/2 & 0 \\ 0 & 1/2 & 1 & 1/2 \\ 0 & 0 & 1/2 & 1 \end{pmatrix}$$

$D_1 = 1$

$D_2 = 1 - 1/4 = 3/4$

$D_3 = -\frac{1}{2}\left(\frac{1}{2}\right) + 1\left(3/4\right) = 2/4 = 1/2$

$$D_4 = -\frac{1}{2}\begin{vmatrix} 1 & 1/2 & 0 \\ 1/2 & 1 & 0 \\ 0 & 1/2 & 1/2 \end{vmatrix} + 1 \cdot \frac{1}{2}$$

$$= -\frac{1}{2}\left(\frac{1}{2}\,\frac{3}{4}\right) + \frac{1}{2} = -\frac{3}{16} + \frac{8}{16} = \frac{5}{16} > 0$$

A pos. defn.

$\Updownarrow$

$H(f) \doteq 2A$ pos. defn.

$$= \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \Rightarrow H(f) \text{ pos. defn for all } (x,y,z,w)$$

$\Rightarrow$ f is convex

Stationary pts: $2A\underline{x} + B^T = 0$ are global mins

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

$$\underbrace{\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}}_{2A} \underbrace{\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}}_{\underline{x}} = \underbrace{\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}}_{-B^T} :$$

Solve for stationary pts using Gaussian elimination

$$\left(\begin{array}{cccc|c} 2 & 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & 2 & -1 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 2 & 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & 2 & -1 \end{array}\right) \xrightarrow{-2} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 0 & -3 & -2 & 0 & -1 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & 2 & -1 \end{array}\right)$$

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

$$\to \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 4 & 3 & -4 \end{array}\right) \begin{matrix} \\ \\ \rfloor -4 \\ \\ \end{matrix} \to \left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & -5 & 0 \end{array}\right)$$

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

$y+2(-1)+1\cdot 0=-1 \quad \underline{y=1}$

$z+2\cdot 0=-1 \quad \underline{z=-1}$

$-5w=0 \quad \underline{w=0}$

Stationary pt: $(x,y,z,w) = \underline{(0,1,-1,0)}$

f convex $\Rightarrow$ $f_{min} = f(0,1,-1,0) = 1^2+(-1)^2-1-1+(-1)-1 = \underline{\underline{-2}}$

(global min. value)

Note: 1) $|A| \neq 0 \Rightarrow$ one stationary pt (one unique sol'n of linear system = FOC)

2) A positive defn. $\Rightarrow$ global min at stationary pt.