7.6 QUADRATIC FORMS + POSITIVE DEFINITE MATRICES

Let $f : \mathbb{R}^n \to \mathbb{R}$, $\quad z = f(\vec{x}) = f(x_1, \ldots, x_n)$

Suppose f is C^2, i.e. $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ $\exists$ + CTS $\forall i,j$. $\forall \vec{x} \in \mathbb{R}^n$

GOAL Find local max/min of f.

CASE $n=2$ (CALCULUS III)

$z = f(x_1, x_2)$

Taylor Expansion of f about $(x_1, x_2) = (a_1, a_2)$ is

$$
\begin{aligned}
f(x_1, x_2) = {} & f(a_1, a_2) + \frac{\partial f}{\partial x_1}(\vec{a})(x_1 - a_1) + \frac{\partial f}{\partial x_2}(x_2 - a_2) \\
& + \frac{\partial^2 f}{\partial x_1^2}(\vec{a})(x_1 - a_1)^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{a})(x_1 - a_1)(x_2 - a_2) \\
& + \frac{\partial^2 f}{\partial x_2^2}(\vec{a})(x_2 - a_2)^2 + O(|\vec{x} - \vec{a}|^3).
\end{aligned}
$$

Write $\nabla f(\vec{a}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(\vec{a}) \\ \dfrac{\partial f}{\partial x_2}(\vec{a}) \end{pmatrix}$ GRADIENT of f.

$$
Hf(\vec{a}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix}\Bigg|_{\vec{x} = \vec{a}}.
$$

$Hf(\vec{a})$ = Hessian of f at $\vec{a}$

$$= \begin{pmatrix} A & B \\ B & C \end{pmatrix} \quad \text{is Symmetric}$$

GENERAL CASE

$$f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x}-\vec{a}) + (\vec{x}-\vec{a})^T Hf(\vec{a}) (\vec{x}-\vec{a})$$
$$+ O(|\vec{x}-\vec{a}|^3)$$

where $[Hf(\vec{a})]_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})$ is Symmetric $n \times n$.

If $\vec{x} = \vec{a}$ is a local max/min then the tangent plane must be horizontal, ie $\nabla f(\vec{a}) = \vec{0}$.

(We say f has a CPt at $\vec{x} = \vec{a}$).

In that case

$$f(\vec{x}) - f(\vec{a}) \simeq \underset{\cancel{Q(\vec{x})}}{Q(\vec{x}-\vec{a})} := (\vec{x}-\vec{a})^T Hf(\vec{a}) (\vec{x}-\vec{a})$$

The behaviour of f near $\overset{\text{cpt}}{\vec{x} = \vec{a}}$ is determined by behaviour of the QUADRATIC FORM Q.

DEF A QUADRATIC FORM is f^n $Q : \mathbb{R}^n \to \mathbb{R}$ given by

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = \sum_{i,j=1}^{n} A_{ij} x_i x_j$$

for some $n \times n$ matrix A.

NOTE The matrix A of Q can always be chosen to be symmetric since for any A

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = \vec{x}^T \underbrace{\left(\frac{A + A^T}{2}\right)}_{\text{SYMMETRIC}} \vec{x}$$

as

$$\underbrace{\vec{x}^T A^T \vec{x}}_{1 \times 1} = \left(\vec{x}^T A^T \vec{x}\right)^T = \vec{x}^T A \vec{x}. \checkmark$$

$\boxed{\text{FROM HERE ON ASSUME } A^T = A}$

~~NOTE~~ GOAL Find local max/min of Q.

CLAIM $\nabla Q(\vec{x}) = A\vec{x}$.

PF $\frac{\partial}{\partial x_i}\left(x^T A x\right) = \frac{\partial \vec{x}^T}{\partial x_i} A\vec{x} + \vec{x}^T A \frac{\partial \vec{x}}{\partial x_i} = \vec{e}_i^{\,T} A\vec{x} + \vec{x}^T A \vec{e}_i$

$= 2\vec{e}_i^{\,T} A\vec{x} = 2(A\vec{x})_i$.

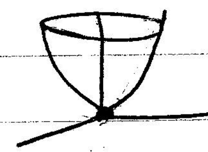
Hence

1. If A is INV Then $\vec{x} = \vec{0}$ is only CPT of Q

So $\vec{x} = \vec{0}$ is local $\begin{bmatrix} \max \\ \min \end{bmatrix}$ iff $\vec{x} = \vec{0}$ is global max/min

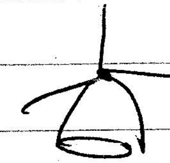
NOTE $Q(\vec{0}) = \vec{0}$.

EG $Q(x_1, x_2) = x_1^2 + x_2^2$



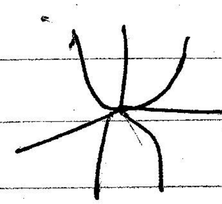
MIN

$Q(x_1, x_2) = -x_1^2 - x_2^2$



MAX

$Q(x_1, x_2) = x_1^2 - x_2^2$



SADDLE PT

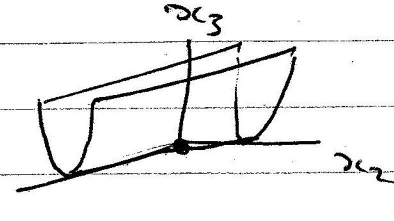
2. If A is SINGULAR Then

CPT SET of $Q = N(A)$

and $Q = 0$ on $N(A)$

EG

$Q(x_1, x_2) = x_2^2$



$N(A) = \{x_2 = 0\}$

Degenerate Case.

DIAGONALIZATION OF Q

$Q(\vec{x}) = \vec{x}^T A \vec{x}$ where $A^T = A$. is real.

We know $A = PDP^T$ where P is orthogonal, D diagonal.

By changing sign of one col of P we can assume $\det(P) = +1$.

Let $R : \mathbb{R}^n \to \mathbb{R}^n$ be the L.T.

$$\vec{y} = R(\vec{x}) = P^T \vec{x}.$$

If $n = 2, 3$, R is a rotation of coord system.

Define $\tilde{Q} = Q \circ R^{-1}$

$$\tilde{Q}(\vec{y}) = Q(R^{-1}\vec{y}) = Q(P\vec{y})$$

$$= \vec{x}^T P D P^T P \vec{y}$$

$$= \vec{y}^T D \vec{y}$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2$$

Suppose A is INV. Then $\lambda_i \neq 0 \ \forall i$ as $\det A = \lambda_1 \dots \lambda_n$.

<u>DEF</u> We say a symmetric $n \times n$ matrix A is <u>positive definite</u> if all evalues of A are positive. Write $A > 0$

APPLICATIONS

① <u>THM</u> [2ND DER TEST]
Let $f : \mathbb{R}^n \to \mathbb{R}$ be C^2 and let $\vec{a} \in \mathbb{R}^n$ be a CPt of f (ie $\nabla f(\vec{a}) = \vec{0}$).

Let $A = Hf(\vec{a})$ be The Hession of f at $\vec{a}$.

Then

① If $A > 0$ Then $\vec{a}$ is a Local Min of f

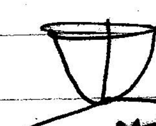
② If $-A > 0$ Then $\vec{a}$ is a Local Max of f

③ If A is INV and has at least one positive and at least one negative evalue Then $\vec{a}$ is a saddle pt of f

④ If A is <u>not</u> INV., no conclusions can be drawn about nature of CPt at $\vec{a}$.

<u>EXS</u> $(n=2)$

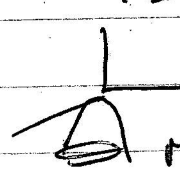
ⓐ $f(x_1, x_2) = x_1^2 + x_2^2$



MIN

$A = I$

ⓑ $f(x_1, x_2) = -x_1^2 - x_2^2$



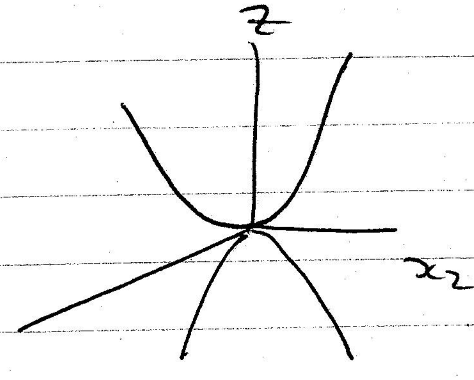
MAX

$A = -I$

(c) $f(x_1, x_2) = x_1^2 - x_2^2$

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

$\vec{0}$ is neither max nor min



SADDLE PT

(II) ELLIPSOIDS

Let $A > 0$, $c > 0$. Then the level surface

$c = Q(\vec{x}) = x^T A \vec{x}$ is an $(n-1)$-DIML ellipsoid with Principal Axes given by evectors of A, eccentricities determined by evalues of A

EG $\boxed{n=2}$

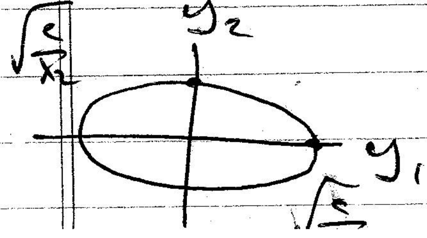
$$A x_1^2 + 2B x_1 x_2 + C x_2^2 = c \quad (*)$$

Suppose $AC - B^2 > 0$, $A > 0$.
Then

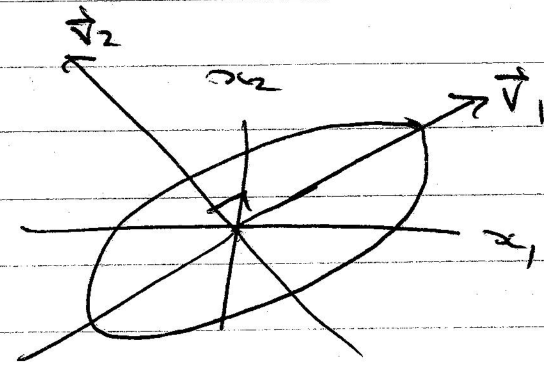
$$\begin{pmatrix} A & B \\ B & C \end{pmatrix} > 0 \quad \text{(See later)}$$

So $(*)$ can be diagonalized to give

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = c$$



ROTATE TO GET:



<u>FORMAL DEF</u> A symmetric matrix A is POS DEF if

$$x^T A x > 0 \quad \forall x \in \mathbb{R}^n \quad x \neq 0$$

<u>THM</u> (TESTS FOR POS DEF)

Let A be symmetric. TFAE

1. $\vec{x}^T A \vec{x} > 0 \quad \forall x \in \mathbb{R}^n \quad x \neq 0,$
2. All evalues of A are positive $(\lambda_i > 0)$
3. $A = B^T B$ for some invertible B
4. Let $A = \left(\begin{array}{c|c} A_{k \times k} & * \\ \hline * & * \end{array}\right)$

   Then $\det(A_k) > 0$ for $k = 1, \ldots, n$
5. Doing row ops of form
   Row $k \to$ Row $k - \alpha$ Row l,

   can reduce A to a row echelon form with all pivots positive.

<u>NOTE</u> <u>Case $n=2$</u> $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ (4): $ad - b^2 > 0$, $a > 0$.
See Math 251 Criterion

PF

$\boxed{① \Rightarrow ③}$ If $Av = \lambda v$ $v \neq 0$ Then

$0 < v^T A v = v^T \lambda v = \lambda \|v\|^2.$

So $\lambda > 0.$

$\boxed{② \Rightarrow ③}$ $A = PDP^T$ $D = \mathrm{Diag}[\lambda_1 \_\_ \lambda_n]$ $\lambda_j > 0.$

Set $D^{1/2} = \mathrm{Diag}\,[\lambda_1^{1/2} \_\_ \lambda_n^{1/2}]$

and $B = (P D^{1/2})^T$

Then UI/ $A = B^T B$, B INV.

$\boxed{③ \Rightarrow ①}$ If $A = B^T B$ Then

$x^T A x = x^T B^T B x = (Bx)^T (Bx) = \|Bx\|^2 > 0$

if $x \neq 0$ since B is INV.

~~ASIDE~~

$\boxed{1 \Rightarrow 4}$

$\boxed{k=n}$ $A_{k\times k} = A.$

$\det(A) = \lambda_1 \ldots \lambda_n > 0$ by $1 \Rightarrow 2$.

$\boxed{\text{OTHER } k}$ Let $\vec{x} = \begin{bmatrix} \vec{x}_k \\ \hline 0 \end{bmatrix}$

Then

$$0 < \vec{x}^T A \vec{x} = \left[\vec{x}_k^T \mid 0 \right] \left[\begin{array}{c|c} A_{k\times k} & * \\ \hline * & * \end{array} \right] \begin{bmatrix} \vec{x}_k \\ \hline 0 \end{bmatrix}$$

$$= \vec{x}_k^T A_{k\times k} \vec{x}_k.$$

NOW apply $1 \Rightarrow 4$ for "$k=n$" to ~~pr~~ conclude

$$\det(A_{k\times k}) > 0.$$

$\boxed{4 \Rightarrow 5}$

Let $d_1 \ldots d_n$ be pivots in RE form of A.
When do specified Row Ops in standard order
find that
$d_1 \ldots d_k$ are pivots in RE form of $A_{k\times k}$.

Since $\det(A_{k\times k}) = d_1 \ldots d_k \; \forall k$ we find

$$d_k = \begin{cases} \det(A_{1\times 1}) & k=1 \\ \det(A_{k\times k}) / \ldots & k>1 \end{cases}$$

$\boxed{⑧ \Rightarrow ①}$

RECALL

Any $n \times n$ A can be converted to row echelon form $\tilde{U}$ by row ops that correspond to mult^n by an elementary matrix.

$$\tilde{U} = EA \qquad E \text{ INV} \qquad \tilde{U}, \text{ T or.}$$

So $$A = L\tilde{U} \qquad L = E^{-1}$$

UT✓ can choose EROs so that E, L are lower Δ^r.

So $$\boxed{A = LDU} \qquad \tilde{U} = DU \qquad D = \text{Diag}(d_1 \_ d_n)$$
$$d_j = \text{PIVOTS}$$

If $A^T = A$ Then $U = L^T$ holds.

So $$A = LDL^T$$

Then
$$\vec{x}^T A \vec{x} = (\vec{x}^T L)\, D (L^T \vec{x})$$
$$= \sum_{j=1}^{n} d_j \left(L^T \vec{x}\right)_j^2 > 0 \quad \text{if } d_j > 0.$$

$\Box$