④ For any symmetric $n\times n$ matrix $\underset{\underset{(a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}}{\parallel}}{A} = A^T$ with $a_{ij}=a_{ji}$,

there is an associated quadratic form

$$Q_A(\bar{x}) \underset{\text{DEF'N}}{:=} \bar{x}^T A\bar{x} \;=\; \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{1\le i<j\le n} 2a_{ij}x_ix_j$$

---

e.g. $n=2$

$$A=\begin{bmatrix} a_{11} & a_{12}\\ a_{12} & a_{22}\end{bmatrix} \text{ has } Q_A\begin{pmatrix}x\\y\end{pmatrix} = [x\ y]\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

$$= [x\ y]\begin{bmatrix} a_{11}x+a_{12}y\\ a_{12}x+a_{22}y\end{bmatrix}$$

$$= a_{11}x^2 + a_{12}xy + a_{12}xy + a_{22}y^2$$

$$= a_{11}x^2 + a_{22}y^2 + 2a_{12}xy$$

And, in fact, every quadratic form $Q(\bar{x}) = \sum_{i=1}^{n} a_ix_i^2 + \sum_{1\le i<j\le n} b_{ij}x_ix_j$

has $Q = Q_A$ for some symmetric A, namely let $a_{ii} := a_i$, $a_{ij} = \frac{1}{2}b_{ij}$, i.e. $A=\begin{bmatrix} a_1 & \frac{b_{12}}{2} & \frac{b_{13}}{2} & \cdots \\ \frac{b_{12}}{2} & a_2 & & \\ \frac{b_{13}}{2} & & \ddots & \\ \vdots & & & a_n\end{bmatrix}$

2/3/2017

⑤ The quadratic Taylor polynomial $P^2_{f,\bar{0}}(\bar{x})$ for $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$

looks like

$$P^2_{f,\bar{0}}(\bar{x}) = f(\bar{0}) + \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\bar{0})\cdot x_i + \sum_{i=1}^{n}\frac{\frac{\partial^2 f}{\partial x_i^2}(\bar{0})}{2!}\cdot x_i^2 + \sum_{1\le i<j\le n}\frac{\partial^2 f}{\partial x_i\partial x_j}(\bar{0})x_ix_j$$

$$= f(\bar{0}) + \underset{\substack{\text{gradient of } f\\ \text{at } \bar{0}}}{\nabla f(\bar{0})^T}\bar{x} + \frac{1}{2}\overbrace{Q_H(\bar{x})}^{\bar{x}^T H\bar{x}} \quad \text{where}$$

$$H = \left[\frac{\partial^2 f}{\partial x_i\partial x_j}(\bar{0})\right]_{\substack{i=1,\dots,n\\ j=1,\dots,n}} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\bar{0}) & \frac{\partial^2 f}{\partial x_1\partial x_2}(\bar{0}) & \cdots \\ \frac{\partial^2 f}{\partial x_1\partial x_2}(\bar{0}) & \ddots & \\ \vdots & & \frac{\partial^2 f}{\partial x_n^2}(\bar{0})\end{bmatrix}$$

is called the Hessian matrix of f at $\bar{x}=\bar{0}$.

(17) What plays the role of $f(x) = f(0) + \underbrace{f'(0)x}_{\text{zero at local extrema}} + \frac{f''(0)}{2}x^2 + \ldots$ having $f''(0) > 0$ or < 0 ? $= 0$

THM 3.5.3 ("Sylvester's law of Inertia")

(i) Every quadratic form $\mathbb{R}^n \xrightarrow{Q} \mathbb{R}$, $\bar{x} \longmapsto Q(\bar{x})$ has an expression

$$Q(\bar{x}) = \alpha_1(\bar{x})^2 + \ldots + \alpha_k(\bar{x})^2 - \alpha_{k+1}(\bar{x})^2 - \ldots - \alpha_{k+l}(\bar{x})^2$$

where the $k+l$ functions $\mathbb{R}^n \xrightarrow{\alpha_i} \mathbb{R}$

are each linear (of the form $\alpha_i(\bar{x}) = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n$
$= [a_{i1} \cdots a_{in}]^T \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$)

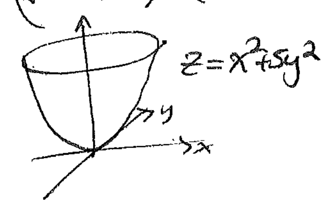
and are linearly independent
in the sense that the $\left\{ \begin{bmatrix} a_{i1} \\ \vdots \\ a_{in} \end{bmatrix} : i = 1, 2, \ldots, k+l \right\}$ are lin. indep.

(ii) Such expressions for Q may not be unique, but the pair (k, l), called the signature of Q, is unique.

EXAMPLES:

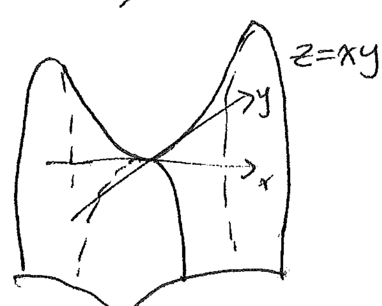
① $Q\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 5y^2 = \frac{1}{2}\left[(x - \sqrt{5}y)^2 + (x + \sqrt{5}y)^2\right] = \left(\frac{x}{\sqrt{2}} - \frac{\sqrt{5}}{\sqrt{2}}y\right)^2 + \left(\frac{x}{\sqrt{2}} + \frac{\sqrt{5}}{\sqrt{2}}y\right)^2$

has signature $(2, 0)$



② $Q\begin{pmatrix} x \\ y \end{pmatrix} = xy = \frac{1}{4}\left[(x+y)^2 - (x-y)^2\right]$
$= \left(\frac{x}{2} + \frac{y}{2}\right)^2 - \left(\frac{x}{2} - \frac{y}{2}\right)^2$

has signature $(1, 1)$



③ $Q\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x + 2y + 3z)^2$

has signature $(1, 0)$

(18)

Proof of THM 3.5.3:

The existence of such an expression for $Q(\bar{x}) = \bar{x}^T A \bar{x}$ with A symmetric follows from the Spectral Theorem, which lets one write $A = P \Lambda P^{-1}$ where P is orthogonal, so $P^{-1} = P^T$

$$= P \Lambda P^T$$

and $\Lambda = \begin{bmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_k & & & & & & \\ & & & -\mu_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -\mu_l & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{bmatrix}$ (k positive, l negative, zeroes) is diagonal

The book gives a different, algorithmic proof via completing the square

Then $Q(\bar{x}) = \bar{x}^T A \bar{x}$

$$= \bar{x}^T P \Lambda P^T \bar{x}$$

$$= (P^T\bar{x})^T \begin{bmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_k & & & & & & \\ & & & \mu_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \mu_l & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{bmatrix} (P^T\bar{x})$$

$$= \lambda_1 \pi_1(\bar{x})^2 + \cdots + \lambda_k \pi_k(\bar{x})^2 - \mu_1 \pi_{k+1}(\bar{x})^2 - \cdots - \mu_l \pi_{k+l}(\bar{x})^2$$

where $\pi_i(\bar{x}) := [i^{\text{th}} \text{ col of } P] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$$\cancel{= \left(\frac{\pi_1(\bar{x})}{\sqrt{\lambda_1}}\right)^2 + \cdots + \left(\frac{\pi_k(\bar{x})}{\sqrt{\lambda_k}}\right)^2 - \left(\frac{\pi_{k+1}(\bar{x})}{\sqrt{\mu_1}}\right)^2 - \cdots - \left(\frac{\pi_{k+l}(\bar{x})}{\sqrt{\mu_l}}\right)^2}$$

$$= \left(\sqrt{\lambda_1}\,\pi_1(\bar{x})\right)^2 + \cdots + \left(\sqrt{\lambda_k}\,\pi_k(\bar{x})\right)^2 - \left(\sqrt{\mu_1}\,\pi_{k+1}(\bar{x})\right)^2 - \cdots - \left(\sqrt{\mu_l}\,\pi_{k+l}(\bar{x})\right)^2$$

$$= \alpha_1(\bar{x})^2 + \cdots + \alpha_k(\bar{x})^2 - \alpha_{k+1}(\bar{x})^2 - \cdots - \alpha_{k+l}(\bar{x})^2.$$

For the uniqueness of (k, l), it helps to introduce these notions ...

DEF'N: (3.5.9 plus a bit more) For a quadratic form $Q(\bar{x})$ on $\mathbb{R}^n$ and a subspace $V \subset \mathbb{R}^n$, say Q is positive definite on V if $Q(\bar{x}) > 0 \ \forall \bar{x} \in V - \{\bar{0}\}$

negative definite on V if $Q(\bar{x}) < 0 \ \forall \bar{x} \in V - \{\bar{0}\}$

(Q is nonnegative definite on V if $Q(\bar{x}) \geq 0 \ \forall \bar{x} \in V$)
or positive semidefinite

(19)

Then the uniqueness of (k,l) will follow from this fact:

PROP 3.5.11: When a quad. form $Q(\bar{x})$ is expressed

$$Q(\bar{x}) = \alpha_1(\bar{x})^2 + \cdots + \alpha_k(\bar{x})^2 - \alpha_{k+1}(\bar{x})^2 - \cdots - \alpha_{k+l}(\bar{x})^2,$$ with $\{\alpha_1, \ldots, \alpha_{k+l}\}$ lin. indep. linear maps $\mathbb{R}^n \to \mathbb{R}$

then $k = \max\{\dim(V) : V$ a subspace of $\mathbb{R}^n$ on which Q is positive definite$\}$

$l = \max\{$ ——— " ——— " ——— negative definite$\}$

proof: It suffices to show the description for k; replacing $Q(\bar{x})$ by $-Q(\bar{x})$ then shows the same for l.

To see $k \geq \max\{\dim(V) : Q$ on $V \underset{\mathbb{R}^n}{\cap}$ is pos. def.$\}$,

assume ~~not maximum~~ $\dim(V) \geq k+1$ and we'll find some $\bar{x} \in V - \{\bar{0}\}$ with $Q(\bar{x}) \leq 0$. Specifically, take any nonzero $\bar{x}$ in the kernel of this linear map

$$\begin{array}{ccc} V & \longrightarrow & \mathbb{R}^k \\ \bar{x} & \longmapsto & \begin{bmatrix} \alpha_1(\bar{x}) \\ \vdots \\ \alpha_k(\bar{x}) \end{bmatrix}, \end{array}$$

which we know exists since $\dim(V) \geq k+1 > \dim(\mathbb{R}^k)$.

Then this nonzero $\bar{x}$ has $Q(\bar{x}) = \alpha_1(\bar{x})^2 + \cdots + \alpha_k(\bar{x})^2 - \alpha_{k+1}(\bar{x})^2 - \cdots - \alpha_{k+l}(\bar{x})^2$

$$= -\alpha_{k+1}(\bar{x})^2 - \cdots - \alpha_{k+l}(\bar{x})^2 \leq 0.$$

To see $k \leq \max\{\dim(V) : Q$ on V is pos. def.$\}$,

we find such a V_0 with $\dim(V_0) = k$ as follows.

Complete $\alpha_1, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_{k+l}$ to a basis $\{\alpha_1, \ldots, \alpha_n\}$ for all linear maps $\mathbb{R}^n \longrightarrow \mathbb{R}$ (Why can we do this?).

Then let $V_0 = \ker\left(\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{\varphi} & \mathbb{R}^{n-k} \\ \bar{x} & \longmapsto & \begin{bmatrix} \alpha_{k+1}(\bar{x}) \\ \vdots \\ \alpha_n(\bar{x}) \end{bmatrix} \end{array} \right)$, which has $\dim(V_0) = k$

(20)

We have Q pos. def. on V_0 since any $\bar{x} \in V_0$ has

$$Q(\bar{x}) = \alpha_1(\bar{x})^2 + \dots + \alpha_k(\bar{x})^2 - \underbrace{\alpha_{k+1}(\bar{x})^2}_{0} - \dots - \underbrace{\alpha_{k+l}(\bar{x})^2}_{0} \geq 0$$

with $Q(\bar{x}) = 0 \iff \alpha_1(\bar{x}) = \dots = \alpha_k(\bar{x}) = 0$

$\iff \alpha_1(\bar{x}) = \dots = \alpha_k(\bar{x}) = \alpha_{k+1}(\bar{x}) = \dots = \alpha_n(\bar{x}) = 0$

since $\bar{x} \in V_0 = \ker \varphi$ $\quad \iff \bar{x} = \bar{0}$

since $\alpha_1, \alpha_2, \dots, \alpha_n$ are a basis for all linear functions $\mathbb{R}^n \to \mathbb{R}$ ∎

2/6/2017 >

This now equips us for...

§3.6 Classifying critical points

THM 3.6.3: Given $\underset{\mathbb{R}^n}{\overset{\text{open}}{U \cap}} \xrightarrow{f} \mathbb{R}$ with f differentiable,

if $\bar{x}_0 \in U$ is an extremum (max or min) for f on U

then $[Df(\bar{x}_0)] = \bar{0}$, i.e. $\frac{\partial f}{\partial x_1}(\bar{x}_0) = \dots = \frac{\partial f}{\partial x_n}(\bar{x}_0) = 0$

$$[Df(\bar{x}_0)] = \left[\frac{\partial f}{\partial x_1}(\bar{x}_0) \dots \frac{\partial f}{\partial x_n}(\bar{x}_0)\right] = \nabla f(\bar{x}_0)^T$$

(in which case we call $\bar{x}_0$ a critical point for f and $f(\bar{x}_0)$ a critical value for f; DEF'N 3.6.4)

proof: We've seen this argument before: assume the min case, and

$$\frac{\partial f}{\partial x_i}(\bar{x}_0) = \lim_{t \to 0} \frac{f(\bar{x}_0 + te_i) - f(\bar{x}_0)}{t}$$ exists since f is diff'ble at $\bar{x}_0$

$$= \begin{cases} \lim_{t \to 0^+} \dfrac{f(\bar{x}_0 + te_i) - f(\bar{x}_0)}{t} & \begin{matrix}\leftarrow \text{nonnegative} \\ \leftarrow \text{positive}\end{matrix} \Rightarrow \geq 0 \\ \lim_{t \to 0^-} \dfrac{f(\bar{x}_0 + t\bar{e}_i) - f(\bar{x}_0)}{t} & \begin{matrix}\leftarrow \text{nonnegative} \\ \leftarrow \text{negative}\end{matrix} \Rightarrow \leq 0. \end{cases}$$

so $\frac{\partial f}{\partial x_i}(\bar{x}_0) = 0$ ∎