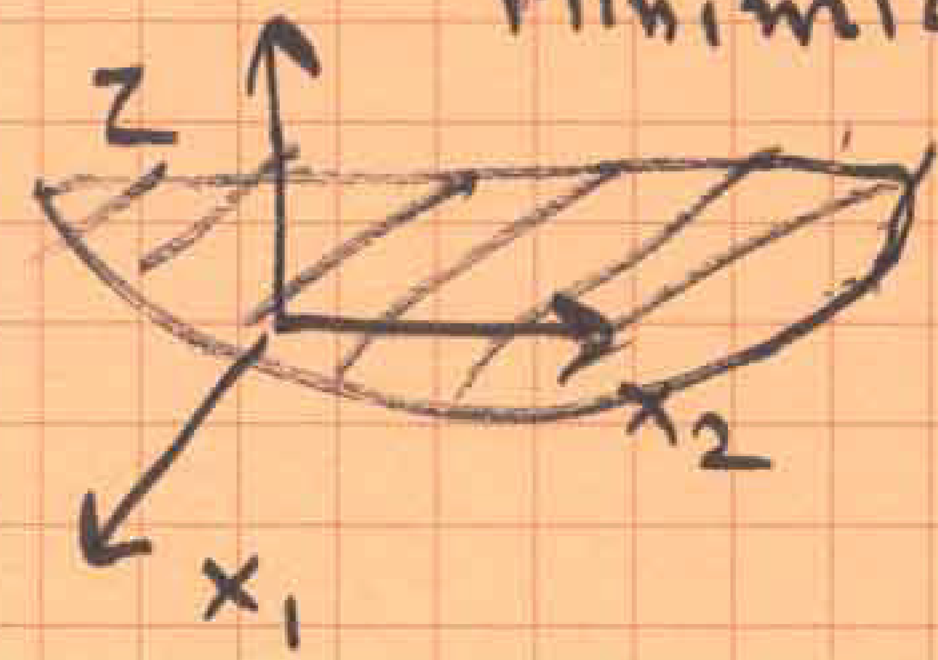
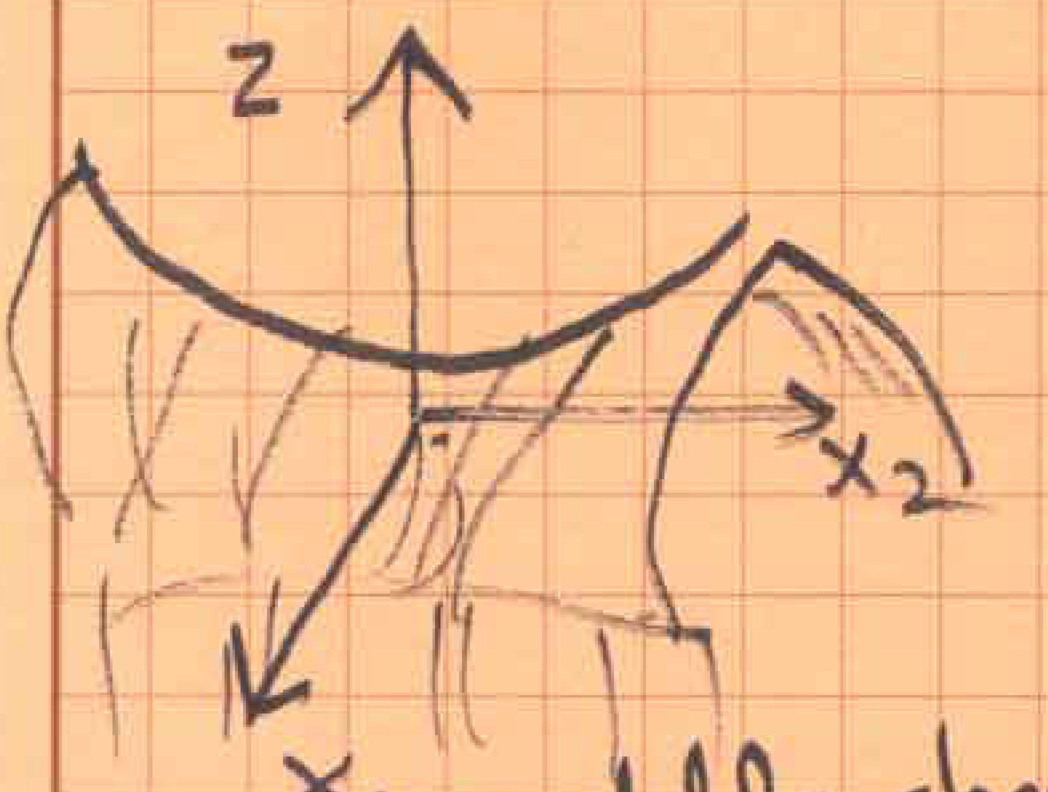
Minimization problems for quadratic surfaces



Ex: $z = p(x_1, x_2) = x_1^2 + 3x_2^2 + 2x_1 - 6x_2 + 1$
$= (x_1+1)^2 + 3(x_2-1)^2 - 3$

minimum value $z = -3$
at $x_1 = -1$, $x_2 = 1$
no max.



saddle shape

Ex: $z = p(x_1, x_2) = -x_1^2 + x_2^2 + x_2$
$= -x_1^2 + (x_2 + \frac{1}{2})^2 - \frac{1}{4}$
no max. or min.

Ex: $p(x_1, x_2) = x_1^2 - x_1 x_2 + x_2^2 + x_1 - x_2 + 1$
$= (x_1 - \frac{1}{2}x_2 + \frac{1}{2})^2 + (x_2 - \frac{1}{4})^2 + \frac{11}{16}$

unique minimum of $\frac{11}{12}$ at $(x_1, x_2) = (-\frac{3}{8}, \frac{1}{4})$
no max.

Ex: $p(x_1, x_2) = x_1^2 + x_1 - 2x_2 + 1$
no max or min

Ex: $p(x_1, x_2) = x_1^2 + 5x_1 - 1$
$= (x_1 + \frac{5}{2})^2 - \frac{29}{4}$

min $-\frac{29}{4}$ at all pts. $(x_1, x_2) = (-\frac{5}{2}, x_2)$
min. not unique.

Where do these problems arise:

(*) $\underset{m\times n}{A}\ \underset{n\times 1}{\underset{\sim}{x}} = \underset{m\times 1}{\underset{\sim}{b}}$

It may be that (*) has no exact soln. $\underset{\sim}{x}$, especially if $m >> n$. (overdetermined system)

Try to find the "best" approximate soln by minimizing $p(\underset{\sim}{x}) = \|A\underset{\sim}{x} - \underset{\sim}{b}\|^2$

Euclidean norm
$= \langle A\underset{\sim}{x} - \underset{\sim}{b}, A\underset{\sim}{x} - \underset{\sim}{b}\rangle$

$\underset{\sim}{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \underset{\sim}{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$

$A = (\underset{\sim}{v}_1, \cdots, \underset{\sim}{v}_n)$

$\therefore p(\underset{\sim}{x}) = \langle x_1\underset{\sim}{v}_1 + \cdots + x_n\underset{\sim}{v}_n - \underset{\sim}{b}, x_1\underset{\sim}{v}_1 + \cdots + x_n\underset{\sim}{v}_n - \underset{\sim}{b}\rangle$
$= \underset{\sim}{x}^T K \underset{\sim}{x} - 2\underset{\sim}{x}^T \underset{\sim}{f} + c$

In particular,

$p(\underset{\sim}{x}) = (A\underset{\sim}{x} - b)^T(A\underset{\sim}{x} - \underset{\sim}{b})$
$= \underline{\underline{x^T A^T A \underset{\sim}{x}}} - 2\underset{\sim}{x}^T A^T \underset{\sim}{b} + \underset{\sim}{b}^T \underset{\sim}{b}$
$= x^T K \underset{\sim}{x} - 2\underset{\sim}{x}^T \underset{\sim}{f} + c$

$\underset{n\times n}{K} = K^T = \underset{n\times m}{A^T}\underset{m\times n}{A}, \quad \underset{\sim}{f} = A^T \underset{\sim}{b}, \quad c = \underset{\sim}{b}^T \underset{\sim}{b}$

To find best approx (least squares approx) must find the minimum of $p(\underset{\sim}{x})$.

Closest point to a subspace:

<u>space</u> R^m, dot product — V_n an n-dim. subspace of R^m, $n \le m$

$\underset{\sim}{b} \in R^m$

Find the point $\underset{\sim}{v}_0 \in V_n$ that is closest to $\underset{\sim}{b}$ in the norm.

$$\text{minimise } \|\underset{\sim}{v} - \underset{\sim}{b}\|^2, \quad \underset{\sim}{v} \in V_n$$

Take basis $\{\underset{\sim}{v}_1, \dots, \underset{\sim}{v}_n\}$ for V_n.

Then $\underset{\sim}{v} = x_1 \underset{\sim}{v}_1 + \dots + x_n \underset{\sim}{v}_n$; unique coords

$$p(\underset{\sim}{x}) = \|\underset{\sim}{v} - \underset{\sim}{b}\|^2 = \langle x_1\underset{\sim}{v}_1 + \dots + x_n \underset{\sim}{v}_n - \underset{\sim}{b},\ x_1 \underset{\sim}{v}_1 + \dots + x_n \underset{\sim}{v}_n - \underset{\sim}{b} \rangle$$

$$= \underset{\sim}{x}^T K \underset{\sim}{x} + \underset{\sim}{x}^T \underset{\sim}{f} + c$$

Same problem!

<u>Problem</u>: Find the minimum of the quadratic form $\boxed{p(\underset{\sim}{x}) = \underset{\sim}{x}^T K \underset{\sim}{x} - 2\underset{\sim}{x}^T \underset{\sim}{f} + c \quad (*)}$

where $\underset{n\times n}{K} = K^T$, $\underset{n\times 1}{\underset{\sim}{x}}$, $\underset{n\times 1}{\underset{\sim}{f}}$, $\underset{1\times 1}{c}$,

if a minimum exists.

Here, $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle = \underset{\sim}{x}^T \underset{\sim}{y}$

<u>Theorem</u> If K is pos. def. then $(*)$ has a <u>unique</u> minimum at $\underset{\sim}{x} = \underset{\sim}{x}^*$ where $\boxed{K\underset{\sim}{x}^* = \underset{\sim}{f}}$. ~~The minimum value is $p(x^*)$ =~~

Note: K pos. def $\Rightarrow$ K nonsingular $\Rightarrow x^* = K^{-1}\underset{\sim}{f}$

What is $p(\underset{\sim}{x}^*)$? $\quad K\underset{\sim}{x}^* = \underset{\sim}{f} \Rightarrow \underset{\sim}{x}^{*T} K = \underset{\sim}{f}^T$

$$p(\underset{\sim}{x}^*) = \underset{\sim}{x}^{*T} K \underset{\sim}{x}^* - 2\underset{\sim}{x}^{*T}\underset{\sim}{f} + c$$
$$= \underline{-\underset{\sim}{x}^{*T} K \underset{\sim}{x}^* + c}$$

__Proof of Theorem__ Look at $p(\underset{\sim}{x}) - p(\underset{\sim}{x}^*)$

$$p(\underset{\sim}{x}) - p(\underset{\sim}{x}^*) = \left(\underset{\sim}{x}^T K \underset{\sim}{x} - 2\underset{\sim}{x}^T \underset{\sim}{f}\right) + \underset{\sim}{x}^{*T} K \underset{\sim}{x}^*$$

(Complete the square!)

$$= (\underset{\sim}{x} - \underset{\sim}{x}^*)^T K (\underset{\sim}{x} - \underset{\sim}{x}^*) + \cancel{2\underset{\sim}{x}^T \underbrace{K\underset{\sim}{x}^*}_{\underset{\sim}{f}}}$$
$$\cancel{-2\underset{\sim}{x}^T\underset{\sim}{f}}$$
$$= (\underset{\sim}{x} - \underset{\sim}{x}^*)^T K (\underset{\sim}{x} - \underset{\sim}{x}^*) \geq 0 \quad \text{because } K \text{ pos. def.}$$
$$= 0 \iff \underset{\sim}{x} = \underset{\sim}{x}^*$$

Q.E.D.

__Theorem__: Suppose K is semi-pos.def.
i.e. $\underset{\sim}{x}^T K \underset{\sim}{x} \geq 0$ for all $\underset{\sim}{x}$, and suppose
$\underset{\sim}{f} \in \text{rng}(K)$, i.e. the eqn. $K\underset{\sim}{x} = \underset{\sim}{f}$
has at least one soln: $\underset{\sim}{x}^*$.
Then $p(\underset{\sim}{x}^*)$ is the unique minimum,
but $\underset{\sim}{x}^*$ is __not unique__.

Proof: Same as above.

In all other cases for K, $\underset{\sim}{f}$
$p(\underset{\sim}{x})$ doesn't have a minimum.

Ex: $n=3$, $K = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 5 & 2 \\ 2 & 2 & 9 \end{pmatrix}$, $\underset{\sim}{f} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$

$\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

$P(\underset{\sim}{x}) = \underset{\sim}{x}^T K \underset{\sim}{x} - 2\underset{\sim}{x}^T \underset{\sim}{f} + 1$

$$\boxed{K \underset{\sim}{x}^* = f = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}}$$

$c = 1$

augmented matrix

$$\left(\begin{array}{ccc:c} 1 & 2 & 2 & 1 \\ 2 & 5 & 2 & 0 \\ 2 & 2 & 9 & -1 \end{array}\right) \to \left(\begin{array}{ccc:c} 1 & 2 & 2 & 1 \\ 0 & 1 & -2 & -2 \\ 0 & -2 & 5 & -3 \end{array}\right) \to \left(\begin{array}{ccc:c} 1 & 2 & 2 & 1 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 1 & -7 \end{array}\right)$$

$$= \left(U \begin{array}{:c} 1 \\ -2 \\ -7 \end{array}\right)$$

$\Rightarrow K$ is pos. def

$U \underset{\sim}{x}^* = \begin{pmatrix} 1 \\ -2 \\ -7 \end{pmatrix}$

$x_3^* = -7$

$x_2^* = 2x_3 - 2 = -16$

$x_1^* = -2x_2 - 2x_3^* + 1 = 47$

$\underset{\sim}{x}^* = \begin{pmatrix} 47 \\ -16 \\ -7 \end{pmatrix}$, $P(\underset{\sim}{x}^*) = x^{*T} K \underset{\sim}{x}^* - 2\underset{\sim}{x}^{*T} \underset{\sim}{f} + 1$

$= -\underset{\sim}{x}^{*T} \underset{\sim}{f} + 1$

$= -(47, -16, -7) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + 1$

$= -(47 + 7) + 1 = -53$

min.

Least squares

$\underset{m\times n}{A}, \underset{m\times 1}{\underset{\sim}{b}}, \underset{n\times 1}{\underset{\sim}{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \langle \underset{\sim}{u}, \underset{\sim}{v} \rangle = \underset{\sim}{u}^T \underset{\sim}{v}$

$\boxed{\text{Assume } \ker(A) = \{\underset{\sim}{0}\}}$

minimise $\| A\underset{\sim}{x} - \underset{\sim}{b} \|^2$

$$\underline{p(\underset{\sim}{x})} = (A\underset{\sim}{x} - \underset{\sim}{b})^T (A\underset{\sim}{x} - \underset{\sim}{b})$$
$$= \underset{\sim}{x}^T A^T A \underset{\sim}{x} - 2\underset{\sim}{x}^T A^T \underset{\sim}{b} + \underset{\sim}{b}^T \underset{\sim}{b}$$
$$= \underset{\sim}{x}^T K \underset{\sim}{x} - 2\underset{\sim}{x}^T \underset{\sim}{f} + c$$

$\underset{n\times n}{K} = A^T A$, $\ker(A) = \{0\} \Rightarrow$ columns of A are lin. ind.

Gram matrix

$A = (\underset{\sim}{a}_1, \dots, \underset{\sim}{a}_n)$

$\therefore K$ is pos. def

$k_{ij} = \underset{\sim}{a}_i^T \underset{\sim}{a}_j$

$$\underset{\sim}{f} = A^T \underset{\sim}{b}, \quad c = \underset{\sim}{b}^T \underset{\sim}{b}, \quad K = A^T A$$

$\therefore p(\underset{\sim}{x}) = \| A\underset{\sim}{x} - \underset{\sim}{b} \|^2$ has a unique minimum $\underset{\sim}{x}^*$ where

$K\underset{\sim}{x}^* = \underset{\sim}{f}$, i.e. $\boxed{A^T A \underset{\sim}{x}^* = A^T \underset{\sim}{b}}$

$$\boxed{\underset{\sim}{x}^* = (A^T A)^{-1} A^T \underset{\sim}{b}}$$

Note: $\underset{m\times n}{A} \underset{\sim}{x}^* = \underset{\sim}{b}$ may not have a soln, but

$\boxed{A^T A \underset{\sim}{x}^* = A^T \underset{\sim}{b}}$ always has a unique soln.