

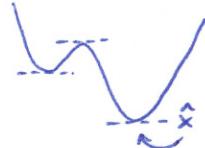
①

## Matrix calculus

Minimizing a smooth function (unconstrained)

$$\hat{x} = \arg \min_{\mathbf{x}} f(\mathbf{x}) \implies f'(\hat{x}) = 0$$

↑  
only necessary,  
not sufficient!

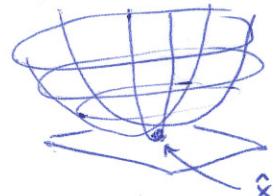


In multiple variables (again smooth + unconstrained)

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad \text{example, } n=2.$$

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} f(\mathbf{x}) \implies \nabla f(\hat{\mathbf{x}}) = 0 \quad (\text{or } \frac{d}{dx} f(x) = 0)$$

↑  
again only  
necessary



$$\nabla f: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

Examples :

$$f(\mathbf{x}) = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (\text{linear function})$$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

written another way ...

$\nabla_{\mathbf{x}} (c^T \mathbf{x}) = c$

(2)

Example #2 : (2-norm)

$$\frac{d}{dx} (\|x\|^2) = \frac{d}{dx} \left( \sum_{i=1}^n x_i^2 \right).$$

$$\frac{\partial}{\partial x_i} \left( \sum_{j=1}^n x_j^2 \right) = 2x_i \Rightarrow \frac{d}{dx} (\|x\|^2) = 2x.$$

Example #3 : (quadratic form)

$$f(x) = x^T Q x \quad (\text{usually assume } Q \text{ is symmetric}) \\ \text{i.e. replace } Q \text{ by } \frac{1}{2}(Q + Q^T).$$

$$\begin{aligned} \frac{d}{dx} x^T Q x &= \frac{d}{dx} \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j \\ \frac{\partial}{\partial x_k} \sum_{j=1}^n \sum_{i=1}^n q_{ij} x_i x_j &= \begin{cases} i=j=k \rightarrow 2q_{kk} x_k \\ i=k, j \neq k \rightarrow q_{kj} x_j \\ j=k, i \neq k \rightarrow q_{ik} x_i \end{cases} \\ &= (Q + Q^T)x. \end{aligned}$$

alternative proof:

$$\frac{d}{dx} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^T \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{d}{dx} \left[ q_{11} x_1^2 + q_{12} x_1 x_2 + q_{13} x_1 x_3 + q_{21} x_2 x_1 + q_{22} x_2^2 + q_{23} x_2 x_3 + q_{31} x_3 x_1 + q_{32} x_3 x_2 + q_{33} x_3^2 \right]$$

$$\begin{aligned} \frac{d}{dx_1} &= 2q_{11}x_1 + q_{12}x_2 + q_{13}x_3 + q_{21}x_2 + q_{31}x_3 \\ &= [q_{11} \ q_{12} \ q_{13}]x + [q_{21} \ q_{31}]x. \end{aligned}$$

$$\frac{d}{dx_2} = [q_{21} \ q_{22} \ q_{23}]x + [q_{12} \ q_{32}]x \quad \left\{ (Q + Q^T)x \right.$$

$$\frac{d}{dx_3} = [q_{31} \ q_{32} \ q_{33}]x + [q_{13} \ q_{23}]x \quad \left. \begin{array}{l} = 2Qx \text{ if } Q = Q^T. \end{array} \right.$$

(3)

## Deriving the normal equation

minimize  $\underset{x}{\|y - Ax\|^2}$  where  $y, A$  are given.

$$\begin{aligned}
 \frac{d}{dx} \|y - Ax\|^2 &= \frac{d}{dx} (y - Ax)^T (y - Ax) \\
 &= \frac{d}{dx} \left( x^T (A^T A) x - \underbrace{y^T A x - x^T A^T y}_{\text{the same!}} + y^T y \right) \\
 &= \frac{d}{dx} \left( x^T (A^T A) x - 2y^T A x \right) \\
 &= 2(A^T A)x - 2A^T y
 \end{aligned}$$

set equal to zero:  $A^T A x = A^T y$  (Normal equations!)

## Quadratic forms and positive-definite matrices

if  $Q \in \mathbb{R}^{n \times n}$  then a function of the form

$$f(x) = x^T Q x$$

is called a quadratic form. Can always assume  $Q$  is symmetric.

(quadratic form for  $Q$  is same as quadratic form for  $\frac{1}{2}(Q+Q^T)$ )

A matrix  $Q \in \mathbb{R}^{n \times n}$  is positive definite if  $x^T Q x > 0$  for all  $x \neq 0$ .

A matrix  $Q \in \mathbb{R}^{n \times n}$  is positive semidefinite if  $x^T Q x \geq 0$  for all  $x$ .

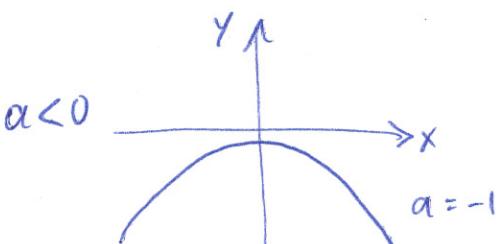
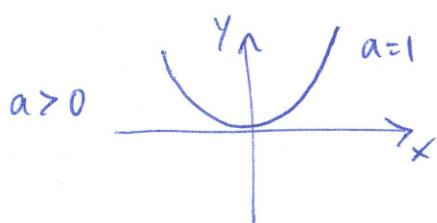
notation:  $Q > 0$  or  $Q \geq 0$

or  $Q > 0$  or  $Q \geq 0$

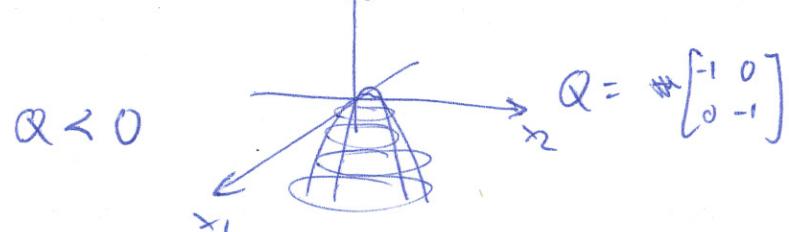
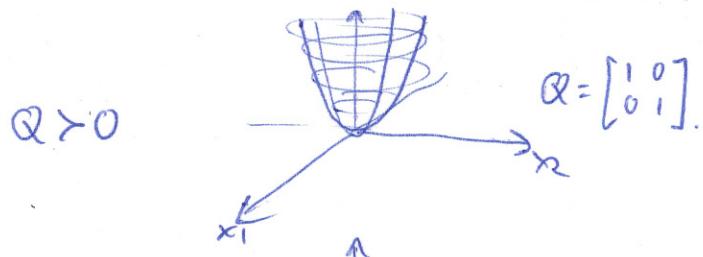
(not same as element-wise positive matrices!)

### graphical

quadratic:  $y = ax^2$

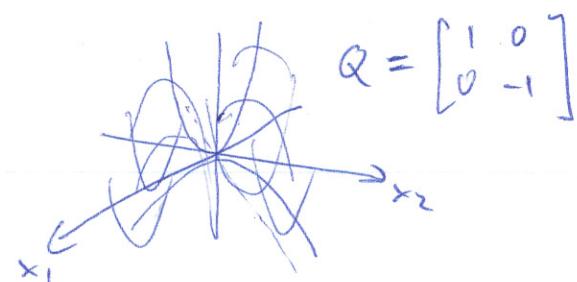


quadratic form:  $y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$



(only possibilities!)

$Q$  indefinite



## Properties of PSD matrices

(5)

- 1) if  $P > 0$  and  $Q > 0$  then  $P+Q > 0$ .
- 2) if  $Q > 0$  then  $\alpha Q > 0$  for all  $\alpha > 0$ .
- 3) if  $P > Q$  and  $Q > R$  then  $P > R$
- 4) if  $P > 0$  then  $-P < 0$ .
- 5) NOT TRUE if  $P > 0$ ,  $Q > 0$  then  $PQ > 0$

## equivalent definition

- (i)  $Q > 0$
  - (ii)  $x^T Q x > 0 \quad \forall x \neq 0,$
  - (iii)  $\lambda_i > 0$  for all eigenvalues of  $Q$ .
- 
- 6) if  $A$  is any matrix,  $A^T A \geq 0$  and  $AA^T \geq 0$ .  
if  $A$  has lin. indep. columns  $\Leftrightarrow A^T A > 0$ .
  - 7) if  $A > 0$  then  $A^{-1}$  exists.
  - 8) if  $A > B > 0$  then  $B^{-1} > A^{-1}$ .