LADS II
1

Transpose: $(A^T)_{ij} = A_{ji}$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

FACT $(AB)^T = \underline{B^T A^T}$.

WATCH OUT: Matlab A' = conjugate tranpose.

Taylor's Theorem for $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}$

cost function.

C^2 - i.e. 2^{nd} partial derivatives exist and are continuous

---

2^{nd} order Taylor expansion at $\vec{x}_0$ is

$$\Phi(\vec{x}_0 + \vec{\Delta x}) = \Phi(x_0) + \nabla\Phi(x_0) \cdot \vec{\Delta x} + \vec{\Delta x}^T H\Phi(x_0) \vec{\Delta x} + \underline{\underline{\text{h.o.t}}}.$$

$$\text{grad } \Phi = \nabla \Phi(x_0) = \left[\frac{\partial \Phi}{\partial x_1}(x_0), \ldots, \frac{\partial \Phi(x_0)}{\partial x_n} \right]$$

nabla, del, grad

$$\text{Hessian of } \Phi = H\Phi(x_0) = \left(\frac{\partial^2 \Phi}{\partial x_i \partial x_j} \right)$$

<u>Since</u> ~~H~~ Φ is C^2 then

$$\frac{\partial^2 \Phi}{\partial x_i \partial x_j} = \frac{\partial^2 \Phi}{\partial x_j \partial x_i} \Rightarrow H^T = H$$

so H is <u>Symmetric</u>

eg// $\Phi(x_1, x_2) = x_1^4 + x_2^4 - 4x_1x_2 + 1$

$$\nabla\Phi(x_1,x_2) = \Big[\underbrace{4x_1^3 - 4x_2}_{\frac{\partial\Phi}{\partial x_1}},\ \underbrace{4x_2^3 - 4x_1}_{\frac{\partial\Phi}{\partial x_2}}\Big]$$

$$H\Phi(x_1,x_2) = \begin{bmatrix} \frac{\partial^2\Phi}{\partial x_1\partial x_1} & \frac{\partial^2\Phi}{\partial x_1\partial x_2} \\ \frac{\partial^2\Phi}{\partial x_2\partial x_1} & \frac{\partial^2\Phi}{\partial x_2\partial x_2} \end{bmatrix}$$

$$= \begin{bmatrix} 12x_1^2 & -4 \\ -4 & 12x_2^2 \end{bmatrix}$$

measures concavity

Introduction to Quadratic Forms

special form

$$\vec{X}^T \begin{bmatrix} a & b \\ b & c \end{bmatrix} \vec{X}$$

$$= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} ax_1 + bx_2, & bx_1 + cx_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= ax_1^2 + bx_2x_1 + bx_1x_2 + cx_2^2$$

$$= ax_1^2 + 2bx_1x_2 + cx_2^2$$

In general, S is symmetric $n \times n$

$$\underline{\vec{x}^T S \vec{x}} = \sum S_{ii} x_i^2 + 2 \sum_{i<j} S_{ij} x_i x_j$$

general quadratic polynomial

What does the graph of

$Q(\vec{x}) = \vec{x}^T S \vec{x}$ look like?

Simplest <u>case</u> $S = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Then $\vec{x}^T S \vec{x}$

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

<u>$n=2$</u> (1) $\lambda_1, \lambda_2 > 0$

graph

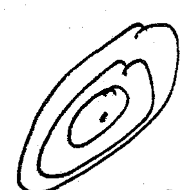
$Q(\vec{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2$

<u>local min.</u>

level sets

(2) $\lambda_1, \lambda_2 < 0$

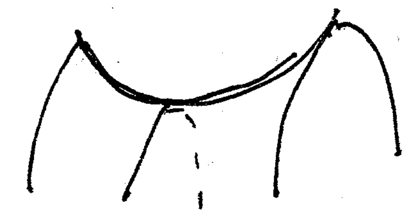
<u>loc MAX</u>



levels sets

(3) $\lambda_1 < 0 \quad \lambda_2 > 0$

$\lambda_1 x_1^2 + \lambda_2 x_2^2$



Saddle point

(4) one or both $\lambda_i = 0$, degenerate case.] no decision

General case of $Q(x) = x^T S x$

real and

S is symmetric $\Longrightarrow$

<u>Spectral Theorem</u> says:

(1) eigenvalues of S are all real.

(2) There is an orthonormal basis of eigenvectors $\vec{u}_1, \ldots, \vec{u}_n$

$\underline{\underline{\text{So}}}$ $U = [\vec{u}_1 \ldots \vec{u}_n]$, <u>$U$ is orthogonal</u> $\Rightarrow$ $U^T = U^{-1}$

and $U^{-1} S U = \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$

$U^T S U = \Lambda$

eigenvalues of S

$Q(x) = x^T S x$ we have

$$U^{-1} S U = \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

then $S = U \Lambda U^{-1} = U \Lambda U^T$

<u>So</u> $x^T S x = x^T U \Lambda U^T x$

new variable $y = U^T x$ then $y^T = x^T U$

$$x^T S x = y^T \Lambda y$$

$\boxed{\Lambda \text{ is diagonal.}}$

the new <u>variable</u>, so back to simplest case.

2nd derivative test for $Q(x) = x^T S x$.

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of S.

(1) all $\lambda_i > 0 \Rightarrow \vec{0}$ is a loc min

(2) all $\lambda_i < 0 \Rightarrow \vec{0}$ is loc MAX.

(3) Some $\lambda_i > 0$, some $\lambda_i < 0 \Rightarrow$ Saddle.

(4) Some $\lambda_i = 0$, no test

Application to <u>optimization</u>

BASIC CALCULUS

Call $\vec{x}_0$ a critical point of Φ if $\nabla\Phi(\vec{x}_0) = 0$.

Theorem: Φ is C^2 then if $\vec{x}_0$ is a local max or min $\Rightarrow$ it's a critical point.

But critical point can be saddle.

Recall. Taylor

$$\Phi(x_0+\Delta x) = \Phi(x_0) + \nabla\Phi(x_0)\,\Delta x + \Delta x^T H\Phi(x_0)\,\Delta x + \text{h.o.t}$$

If x_0 is critical, then $\nabla\Phi(x_0) = 0$

$$\Phi(x_0+\Delta x) \approx \underbrace{\Phi(x_0)}_{\text{Translates up or down}} + \underbrace{\Delta x^T H\Phi(x_0)\,\Delta x}_{\text{quadratic form}}$$

2^{nd} derivative test

If $\nabla\Phi(x_0) = 0$ Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $H\Phi(x_0)$

(1) All $\lambda_i > 0 \Rightarrow x_0$ is a loc min

(2) All $\lambda_i < 0 \Rightarrow x_0$ is loc MAX

(3) Some $\lambda_i < 0$ some $\lambda_i > 0$ $\Rightarrow$ loc saddle

(4) Some $\lambda_i = 0 \Rightarrow$ no test

eg 11 $\Phi(x_1, x_2)$ as above

$$\nabla\Phi(0,0) = 0$$

$$H\Phi(0,0) = \begin{bmatrix} 0 & -4 \\ -4 & 0 \end{bmatrix}$$

$$\begin{vmatrix} -\lambda & -4 \\ -4 & -\lambda \end{vmatrix} = \lambda^2 - 16 \qquad \lambda = \pm 4$$

Saddle

Detailed example next lecture.