Lecture 4

January 30, 2019

Anders Sundheim
asundheim@wisc.edu

Characterization of Critical Points 2

Recall

Let A be a $n \times n$ real, symmetric matrix, and $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$

For $h \in \mathbb{R}^n$, denote by

$$Q(h) = h^T A h = \sum_{i,j=1}^{n} a_{ij} h_i h_j$$

0.1 Lemma

$$\boxed{Q(H) > 0 \text{ for all } h \neq \vec{0} \iff \text{all eigenvalues of } A \text{ are positive}}$$

Proof. Since A is a real, symmetric matrix, we have that A is diagonalizable, and we can write

$$\begin{cases} A = CDC^T \\ D \text{ is a diagonalizable matrix, } D = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ CC^T = I, \text{ that is, } C^T = C^{-1} \end{cases}$$

Then, eigenvalues of A are $\lambda_1, \lambda_2, \dots, \lambda_n$
we note then that:

$$Q(h) = h^T A h = h^T C D C^T h = (h^T C) D (C^T h)$$

Set $x = C^T h \in \mathbb{R}^n \Rightarrow x^T = (C^T h)^T = h^T C$. Then

$$Q(h) = x^T D x = (x_1, \dots, x_n) \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boxed{\sum_{i=1}^{n} \lambda_i x_i^2}$$

The last identity is very important as it confirms what we need

- Firstly, if $\lambda_1, \dots, \lambda_n > 0$, then for $h \neq \vec{0}$, $x = C^T h \neq \vec{0}$, and so $Q(h) = \sum_{i=1}^{n} \lambda_i x_i^2 > 0$
- Secondly, in order for $Q(h) > 0$ for all $h \neq \vec{0}$, we need $\sum_{i=1}^{n} \lambda_i x_i^2 > 0$ for all $x \neq \vec{0}$.

Clearly we yield that $\lambda_1, \dots, \lambda_n > 0$ $\square$

0.2 Second derivative test for critical points of functions of two variables

$$\begin{cases} \text{Let } B(y,r) \subset \mathbb{R}^2, \text{ and } f : B(y,r) \to \mathbb{R} \\ \text{Assume that } \frac{d^2 f}{d_{x_i} d_{x_j}} \text{ exists and continuous in } B(y,r) \text{ for all } 1 \leq i,j \leq 2 \\ \text{Assume further that } f \text{ has a critical point at } y \text{, that is, } \boxed{\nabla f(y) = \vec{0}} \end{cases}$$

Our Hessian at y is:

$$\boxed{H(y) = \begin{pmatrix} f_{x_1 x_1}(y) & f_{x_1 x_2}(y) \\ f_{x_2 x_1}(y) & f_{x_2 x_2}(y) \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix}},$$

where we set $A = f_{x_1 x_1}(y)$, $B = f_{x_1 x_2}(y)$, $C = f_{x_2 x_2}(y)$
To find eigenvalues of $H(y)$, we solve

$$\det(\lambda I - H(y)) = \det \begin{pmatrix} \lambda - A & -B \\ -B & \lambda - C \end{pmatrix} = (\lambda - A)(\lambda - C) - B^2 = 0$$

$$\iff \boxed{\lambda^2 - (A+C)\lambda + AC - B^2 = 0}$$

Let λ_1, λ_2 be the two roots of the above quadratic equation.

$$\text{Then (by Vieta's theorem)} \begin{cases} \lambda_1 + \lambda_2 = A + C \\ \lambda_1 \cdot \lambda_2 = AC - B^2 = \Delta = \det\ H(y) \end{cases} \tag{1}$$

Then we have the following conclusions

1 If $\boxed{\Delta < 0}$, then λ_1, λ_2 have opposite signs $\Rightarrow$ $\boxed{\text{y is a saddle point}}$

2 If $\boxed{\Delta > 0}$, then λ_1, λ_2 have the same signs.

We now just need to determine their signs

Note that

$$\Delta = AC - B^2 > 0 \Rightarrow AC > B^2 \geq 0 \Rightarrow \boxed{AC > 0}$$

$$2.1\ \boxed{\text{If } A > 0} \Rightarrow C > 0 \Rightarrow A + C > 0 \Rightarrow \lambda_1, lambda_2 > 0$$

which gives that $\boxed{\text{y is a local min}}$

$$2.2\ \boxed{\text{If } A < 0}\ \text{Same logic gives}\ \boxed{\text{y is a local max}}$$