Lecture 3

January 28, 2019

Anders Sundheim
asundheim@wisc.edu

1 Characterization of Critical Points

1.1 Critical Points in n dimensions

Recall

$\varphi(t) = f(y + te)$

$\varphi'(t) = f'(y + te, e) = \nabla f(y + te)e$

$$\begin{aligned}
\varphi''(t) &= f''(y + te, e, e) \\
&= \frac{d}{dt}\left(\sum_{i=1}^{n} f_{x_i}(y + te)e_i\right) \\
&= \sum_{i,j=1}^{n} f_{x_i x_j}(y + te)e_i e_j \\
&= eH(y + te)e^T \\
&= \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \begin{pmatrix} f_{x_1 x_1} & \dots & f_{x_1 x_n} \\ \vdots & & \vdots \\ f_{x_n x_1} & \dots & f_{x_n x_n} \end{pmatrix}
\end{aligned}$$

1.2 Hessian

The Hessian of f at x is denoted by

$$H(x) = \begin{pmatrix} f_{x_1x_1}(x) & \dots & f_{x_1x_n}(x) \\ \vdots & & \vdots \\ f_{x_nx_1}(x) & \dots & f_{x_nx_n}(x) \end{pmatrix}$$

$H(x)$ is a $n \times n$ matrix. Since we assume that
$f_{x_ix_j}(x)$ is continuous in $B(y, r)$, we have
$f_{x_ix_j}(x) = f_{x_jx_i}(x) \forall x, i, j$
$\Rightarrow H(x)$ is a real, symmetric matrix
Key Point: This matrix is diagonalizable

1.3 Theorem 1: Taylor expansion of f

Assume f as previous
Then for any $h \in \mathbb{R}^n$ with $|h| < r$

$$f(y+h) = f(y) + \nabla f(y)h + \frac{1}{2}h^T H(y)h + \omega(h)|h|^2$$

where $\lim_{h \to 0} \omega(h) = 0$

1.4 Characterize Critical points

Assume f has a critical point at y, that is $\nabla f(y) = \vec{0}$
$\Rightarrow$ then for $|h| < r$

$$f(y+h) = f(y) + \frac{1}{2}h^T H(y)h + \omega(h)|h|^2$$

$$\begin{aligned} \text{Define } Q(h) &= h^T H(y)h \\ &= \sum_{i,j=1}^{n} f_{x_ix_j}(y) h_i h_j \end{aligned}$$

(If $n = 1$, $Q(h) = f''(y)h^2$)

1.5 Theorem 2: Characterization of critical points 1

1 If $Q(h) > 0$ for $h \neq 0$, then f has a local min at y
2 If $Q(h) < 0$ for $h \neq 0$, then f has a local max at y
3 If we can find $h_1, h_2 \neq 0$ such that $Q(h_1) > 0 > Q(h_2)$, then y is a saddle point

1.6 Theorem 3: Characterization of critical points 2

1 $\iff$ all eigenvalues of $H(y)$ are positive
2 $\iff$ all eigenvalues of $H(y)$ are negative
3 $\iff$ $H(y)$ has at least one positive and negative eigenvalue
If some eigenvalues are 0, it's typically inconclusive

1.7 Examples

1.

$$f : \mathbb{R}^2 \to \mathbb{R}$$

$$(x, y) \mapsto f(x, y) = x^2 + 2x + y^2 - 4y$$

Question: Characterize all critical points of f

Solution: Find all critical points by solving

$$\nabla f(x, y) = \vec{0} = (0, 0)$$

$$(2x + 2, 2y - 4) = (0, 0) \Rightarrow (x, y) = (-1, 2)$$

$$H(-1, 2) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$\Rightarrow$ all eigenvalues positive $\Rightarrow (-1, 2)$ is a local min

2.

$$g : \mathbb{R}^2 \to \mathbb{R}$$

$$(x, y) \mapsto g(x, y) = xy$$

$$\nabla g(x, y) = (y, x) = \vec{0}$$

$$(y, x) = (0, 0) \Rightarrow \text{ one critical point at } (0, 0)$$

$$H(0, 0) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{aligned} \text{Eigenvalues} &= \det(\lambda I - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}) \\ &= \det \begin{pmatrix} \lambda & -1 \\ -1 & \lambda \end{pmatrix} \\ &= \lambda^2 - 1 = 0 \\ &= \lambda^2 = 1 \\ &= \lambda = 1, -1 \end{aligned}$$

$\Rightarrow$ One positive, one negative eigenvalue $\Rightarrow$ $(0, 0)$ is a saddle point