COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Math Methods for Political Science

Lecture 11: Multivariate Calculus II

Thibault Vatter <tv2233@columbia.edu>
Department of Statistics, Columbia University

October 30, 2017

Partial derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$ with $f(\mathbf{z}) \equiv f(x_1, \cdots, x_n)$ for $\mathbf{z} = (x_1, \cdots, x_n)$.

Definition 1 (Partial derivative)

The **k-th partial derivative of f at $\mathbf{z}$** is defined as

$$\frac{\partial f(x_1, \cdots, x_n)}{\partial x_k} = \lim_{h \to 0} \frac{f(x_1, \cdots, x_k + h, \cdots, x_n) - f(x_1, \cdots, x_k, \cdots, x_n)}{h}$$

- Interpretation: how a change in the k-th variable affects f while holding the other variables fixed.
- Notation: sometimes, we use $\partial f / \partial x_k$ or even $\partial_k f$.

Example 1 (Partial derivative)

- Let $\mathbf{z} = (x_1, x_2) \in \mathbb{R}^2$ and $f(x_1, x_2) = x_1^3 + x_2 - 1$.
- Let $\mathbf{z} = (x, y, w) \in \mathbb{R}^3$ and $f(x, y, w) = xyw + x^2 w - y^3$.

Second-order/cross partial derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$ with $f(\mathbf{z}) \equiv f(x_1, \cdots, x_n)$ for $\mathbf{z} = (x_1, \cdots, x_n)$.

Definition 2 (Second-order/cross partial derivative)

If $1 \leq i, j \leq n$, then

- $\partial_{i,j} f = \frac{\partial^2 f}{\partial x_i \partial x_j}$ **is the cross-partial derivative,**
- $\partial_{i,i} f = \frac{\partial^2 f}{\partial x_i \partial x_i}$ **is the second-order partial derivative.**

Intuition:

- Cross: how the effect of x_j on f changes as x_i changes.
- Second-order:
  - the first-order derivative of the first-order derivative,
  - how the effect of x_i on f changes as x_i itself changes.

Note that

$$\partial_{i,j} f = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \partial_{j,i} f.$$

Example 2 (Second-order/cross partial derivative)

Suppose a government's probability of re-election depends on the state of the economy, x, and the prime minister's personal popularity, y. Let the probability be denoted by $p(x, y) = xy$.

1. How does probability of re-election change if the economy improves slightly?
2. How does the probability of re-election change if the prime minister's personal popularity increases slightly?
3. How does the effect of the economy change if the prime minister's personal popularity increases slightly?
4. How does the effect of the prime minister's personal popularity change if the economy improves slightly?
5. What happens when $p(x, y) = x(1 - y) + y(1 - x)$. Which theoretical do you find more plausible?

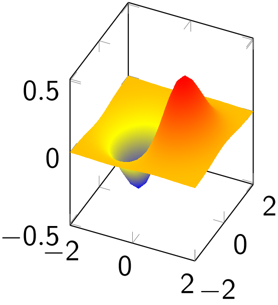
Gradient vector

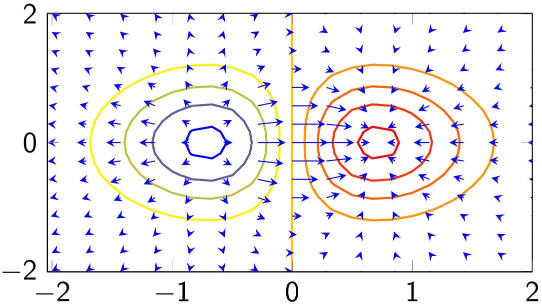
Definition 3 (Gradient vector)

The **gradient is the vector of partial derivatives of f at $\mathbf{z}$**:

$$\nabla f(\mathbf{z}) = \begin{bmatrix} \partial_1 f(\mathbf{z}) \\ \vdots \\ \partial_n f(\mathbf{z}) \end{bmatrix}.$$

Let $f(x, y) = xe^{-x^2-y^2}$.





Local extrema and the gradient

Definition 4 (Local extrema)

If $\exists \delta > 0$ s.t., $\forall \mathbf{z} \in (\mathbf{z}_0 - \delta, \mathbf{z}_0 + \delta) \cap D_f$,

1. $f(\mathbf{z}) \geq f(\mathbf{z}_0)$, then $\mathbf{z}_0$ is a **local minimum**,
2. $f(\mathbf{z}) \leq f(\mathbf{z}_0)$, then $\mathbf{z}_0$ is a **local maximum**.

If 1. or 2. is true, then x_0 is a **local extremum**.

Definition 5 (Stationary point)

$\mathbf{z}_0$ is a **stationary point** if $\nabla f(\mathbf{z}_0) = 0$.

Theorem 1 (Gradient at local extrema)

If f is differentiable at a local extremum $\mathbf{z}_0$, then $\nabla f(\mathbf{z}_0) = 0$.

Example 2 (Local extrema and the gradient)

Consider $f(x, y) = x^2 + y^2$.

Hessian matrix

Definition 6 (Hessian matrix)

The hessian is **the matrix of cross and second-order partial derivatives** of f at $\mathbf{z}$:

$$H(\mathbf{z}) = \begin{bmatrix} \partial_{1,1} f(\mathbf{z}) & \partial_{1,2} f(\mathbf{z}) & \cdots & \partial_{1,n} f(\mathbf{z}) \\ \partial_{2,1} f(\mathbf{z}) & \partial_{2,2} f(\mathbf{z}) & \cdots & \partial_{2,n} f(\mathbf{z}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{n,1} f(\mathbf{z}) & \partial_{n,2} f(\mathbf{z}) & \cdots & \partial_{n,n} f(\mathbf{z}) \end{bmatrix}.$$

Example 3 (Hessian matrix)

- $f(x, y) = x^2 y$
- $f(x, y, z) = x^2/2 + y^2/2 + z^2/2$

Note: $\partial_{i,j} f = \partial_{j,i} f \implies H$ **is symmetric.**

Definiteness

Let H be an $n \times n$ symmetric matrix.

Definition 7 (Definiteness)

H is

1. positive definite if $\mathbf{x}^\top H\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$,
2. positive semidefinite if $\mathbf{x}^\top H\mathbf{x} \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$,
3. negative definite if $\mathbf{x}^\top H\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$,
4. negative semidefinite if $\mathbf{x}^\top H\mathbf{x} \leq 0$ for all $\mathbf{x} \neq \mathbf{0}$,

and it is indefinite otherwise.

Example 4 (Definiteness)

- the identity matrix
- $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Testing for definiteness

Let H be an $n \times n$ symmetric matrix.

Definition 8 (Principal submatrix and minor)

- The **k-th order principal submatrix** of H, H_k, is obtained by deleting the *last* $n-k$ columns and rows.
- Its determinant, $\mid H_k \mid$, is the **k-th order principal minor**.

Example 5 (Principal submatrix and minor)

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \implies H_1 = H_{11},\, H_2 = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix},\, H_3 = H$$

Theorem 2 (Testing for definiteness)

- *H is positive definite* $\mid H_k \mid > 0 \ \forall k$.
- *H is negative definite* $\iff (-1)^k \mid H_k \mid > 0 \ \forall k$.
- *H is indefinite otherwise.*

For semi-definiteness, replace $>$ by $\geq$.

Sufficient condition for extrema

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and H its hessian matrix.

Theorem 3 (Sufficient condition for extrema)

If $\mathbf{z}_0$ is a stationary point, then

1. $H(\mathbf{z}_0)$ *positive definite* $\implies$ $\mathbf{z}_0$ *is a local minimum,*
2. $H(\mathbf{z}_0)$ *negative definite* $\implies$ $\mathbf{z}_0$ *is a local maximum.*

Example 6 (Finding local extrema)

- $f(x, y) = x^2 + y^2$
- $f(x, y) = xe^{-x^2-y^2}$
- $f(x_1, x_2, x_3) = x_1^2 + 3x_2^2 - 3x_1x_2 + 4x_2x_3 + 6x_3^2$