# Theory sheet 9

## Function of many variables

Functions taking $\mathbb{R}^n$ to $\mathbb{R}$ (vectors to numbers) are called *functions of many variables.* Notation:

$$f : \mathbb{R}^n \to \mathbb{R} : (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n).$$

Instead of $f(x_1, \ldots, x_n)$ we also write $f(\boldsymbol{x})$. For example, functions of two variables $f : \mathbb{R}^2 \to \mathbb{R}$ assign a number (output) to each a point on a plane (input).

Some functions are only defined on some domain $D \subset \mathbb{R}^n$. In this case we write

$$f : D \to \mathbb{R}.$$

We have already seen two examples in this class:

- Linear forms $f(\boldsymbol{x}) = \boldsymbol{a}^\top \boldsymbol{x}$, where $\boldsymbol{a} \in M_{n,1}$ is a column vector.
- Quadratic forms $f(\boldsymbol{x}) = \boldsymbol{x}^\top A \boldsymbol{x}$, where $A \in M_{n,n}$ is a symmetric square matrix.

Here's an example from economics: the so-called Cobb-Douglas function $f(K, L)$ takes two positive numbers $K$ (capital) and $L$ (labour) as inputs and outputs the total production:

$$f(K, L) = cK^a L^b,$$

where $c, a, b$ are some parameters.

A simple generalization of the notion of quadratic form is the quadratic function:

$$f : \mathbb{R}^n \to \mathbb{R} : f(\boldsymbol{x}) = \boldsymbol{x}^\top A \boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c,$$

where $A \in M_{n,n}$, $\boldsymbol{b} \in M_{n,1}$ and $c \in \mathbb{R}$.

As we discussed before, a function may be represented by its graph. Here's a reminder of its definition:

$$\text{Graph}(f) = \{(\boldsymbol{x}, y) \in \mathbb{R}^{n+1} : y = f(\boldsymbol{x})\}.$$

Drawing graphs is fine in $n = 1$ and $n = 2$, but as $n$ increases graphs become less helpful.

A very important concept associated with functions $f : \mathbb{R}^n \to \mathbb{R}$ is their level sets, which we have also discussed before. Recall that if $c \in \mathbb{R}$ is some given level, the corresponding level set is the set of points where $f$ assumes this value:

$$\{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) = c\}.$$

Recall the level sets (of height function) on topographic maps!

# Partial derivatives

**Definition 1.** *Partial derivative of* $f : \mathbb{R}^n \to \mathbb{R}$ *with respect to* $x_j$, $j$ *fixed is the following limit:*

$$\frac{\partial f}{\partial x_j} := \lim_{\Delta \to 0} \frac{f(x_1, \ldots, x_j + \Delta, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)}{\Delta}$$

*(provided that it exists). Alternative notation:* $f'_{x_j}$.

In other words, it's just the derivative with respect to $x_j$ with other variables fixed.

Interpretation: the precise definition has $\lim_{\Delta \to 0}$, but we can use $f'_{x_j}(\boldsymbol{x})$ to approximate $f$ at shifted point if $\Delta$ is small:

$$f'_{x_j}(\boldsymbol{x}) \approx \frac{f(x_1, \ldots, x_j + \Delta, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)}{\Delta}$$

$$\Longrightarrow \; f(x_1, \ldots, x_j + \Delta, \ldots, x_n) \approx f(\boldsymbol{x}) + f'_{x_j}(\boldsymbol{x})\,\Delta.$$

Depending on the function and smallness of $\Delta$, this approximation may be accurate or not. If $\Delta$ is not small, this approximation does not make any sense!

Interpretation 2: $f'_{x_j}(\boldsymbol{x})$ is the slope of $f$ at point $\boldsymbol{x}$ in the direction of $x_j$.

## Example

Let

$$f(x_1, x_2) = 2x_1^2 x_2^4 - 5x_1 x_2^3 + 16.$$

Then

$$f'_{x_1}(x_1, x_2) = 4x_1 x_2^4 - 5x_2^3, \quad f'_{x_2}(x_1, x_2) = 8x_1^2 x_2^3 - 15x_1 x_2^2.$$

Let us find the value of these derivatives at $(1, 2)$:

$$f'_{x_1}(1, 2) = 64 - 40 = 24, \quad f'_{x_2}(1, 2) = 64 - 60 = 4.$$

Since both numbers are positive, the function $f$ increases in both variables at $(1, 2)$. This also gives us an approximation of $f(1 + \Delta, 2)$ and $f(1, 2 + \Delta)$ for small $\Delta$:

$$f(1 + \Delta, 2) \approx f(1, 2) + f'_{x_1}(1, 2)\,\Delta = f(1, 2) + 24\,\Delta,$$

$$f(1, 2 + \Delta) \approx f(1, 2) + f'_{x_2}(1, 2)\,\Delta = f(1, 2) + 4\,\Delta.$$

# Elasticity

**Definition 2.** *Elasticity of* $y = f(\boldsymbol{x})$ *with respect to* $x_j$ *is the following limit:*

$$E_{x_j}(y) = \lim_{\Delta \to 0} \frac{\Delta y}{y} \Big/ \frac{\Delta x_j}{x_j},$$

*where* $\Delta y = f(x_1, \ldots, x_j + \Delta, \ldots, x_n) - f(x_1, \ldots, x_j, \ldots, x_n)$ *and* $\Delta x_j = (x_j + \Delta) - x_j = \Delta$.

Elasticity $E_{x_j}(y)$ may be expressed in terms of the partial derivative as follows:

$$E_{x_j}(y) = \frac{x_j}{y} \lim_{\Delta \to \infty} \frac{\Delta y}{\Delta} = \frac{x_j}{y} \frac{\partial y}{\partial x_j}.$$

Example: $y = f(x_1, x_2) = x_1 x_2 e^{x_1 + x_2}$ , then

$$\frac{\partial y}{\partial x_1} = x_2 e^{x_1+x_2} + x_1 x_2 e^{x_1+x_2} = x_2(1 + x_1)e^{x_1+x_2}.$$

Hence,

$$E_{x_1}(y) = \frac{x_1}{y} \frac{\partial y}{\partial x_1} = \frac{x_1}{x_1 x_2 e^{x_1+x_2}} \cdot x_2(1 + x_1)e^{x_1+x_2} = 1 + x_1.$$

**Remark 1.** *We could have arrived at the same solution easier if we noticed that*

$$\frac{1}{y} \frac{\partial y}{\partial x_1} = \frac{\partial}{\partial x_1} \ln y.$$

*This is easier because* ln *takes product into sum:*

$$\ln y = \ln x_1 + \ln x_2 + x_1 + x_2 \implies \frac{\partial}{\partial x_1} \ln y = \frac{1}{x_1} + 1.$$

*It remains to multiply by* $x_1$ *:*

$$E_{x_1}(y) = x_1 \frac{\partial}{\partial x_1} \ln y = x_1 \left( \frac{1}{x_1} + 1 \right) = 1 + x_1.$$

*This trick is known as the* <u>*logarithmic derivative*</u> *and it is frequently useful for differentiating functions defined as* <u>*products*</u> *of simpler terms.*

Another example: Cobb-Douglas function with $b = 1 - a$

$$\begin{aligned} Q = cK^a L^{1-a} \implies E_K(Q) &= K \frac{\partial}{\partial K} \ln Q \\ &= K \frac{\partial}{\partial K} (\ln c + a \ln K + (1 - a) \ln L) \\ &= K \cdot a \cdot \frac{1}{K} = a. \end{aligned}$$

and similarly

$$\begin{aligned} E_L(Q) &= L \frac{\partial}{\partial L} \ln Q \\ &= L \frac{\partial}{\partial L} (\ln c + a \ln K + (1 - a) \ln L) \\ &= L \cdot (1 - a) \cdot \frac{1}{L} = 1 - a. \end{aligned}$$

# Total differential

**Definition 3.** *Total differential or first order differential of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is the following formal object:*

$$df(\boldsymbol{x}) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \, dx_j.$$

You should think about $df$ as a function of $\boldsymbol{x}$ and of formal increments $dx_j$, $j = 1, \ldots, n$. Here $dx_j$ is a formal variable, instead of which we plug some specific $\Delta x_j$ to compute the approximation

$$\Delta f(\boldsymbol{x}) \approx \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \, \Delta x_j.$$

Then $f(\boldsymbol{x} + \boldsymbol{\Delta}) \approx f(\boldsymbol{x}) + \Delta f(\boldsymbol{x})$.

For example, if $f(x_1, x_2) = 2x_1^2 x_2^4 - 5x_1 x_2^3 + 16$, then

$$df(\boldsymbol{x}) = (4x_1 x_2^4 - 5x_2^3) \, dx + (8x_1^2 x_2^3 - 15x_1 x_2^2) \, dy.$$

Taking $\Delta x_1 = 0.02$ and $\Delta x_2 = 0.03$ at $(1, 2)$, we obtain

$$\Delta f(1, 2) = 24 \cdot \Delta x_1 + 4 \cdot \Delta x_2 = 24 \cdot 0.02 + 4 \cdot 0.03 = 0.6.$$

# Gradient

**Definition 4.** *The gradient of* $f : \mathbb{R}^n \to \mathbb{R}$ *is the column vector of partial derivatives:*

$$\operatorname{grad} f(\boldsymbol{x}) = \begin{pmatrix} f'_{x_1} \\ f'_{x_2} \\ \vdots \\ f'_{x_n} \end{pmatrix}.$$

We can rewrite the total differential as

$$df(\boldsymbol{x}) = (\operatorname{grad} f(\boldsymbol{x}))^\top d\boldsymbol{x}.$$

Note that if $\operatorname{grad} f(\boldsymbol{x})$ is orthogonal to a given increment vector $\Delta \boldsymbol{x}$, then

$$\Delta f(\boldsymbol{x}) \approx (\operatorname{grad} f(\boldsymbol{x}))^\top \Delta \boldsymbol{x} = 0.$$

This means that $f$ changes slower than linearly in the direction of $\boldsymbol{\Delta}$.

**Remark 2.** *Gradient is always orthogonal to the level sets. We have discussed this before with linear forms!*

## Higher partial derivatives

Similarly to how we defined $f'_{x_j}$ , we can define partial derivatives of second order derivatives:

$$f''_{x_j,x_i}(\boldsymbol{x}) \coloneqq \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(\boldsymbol{x}).$$

Do we need to keep track of the order in which we compute them? Luckily, no. For *nice* functions partial derivatives commute:

$$f''_{x_j,x_i}(\boldsymbol{x}) = f''_{x_i,x_j}(\boldsymbol{x})$$

or in other notation

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

We can now go further and define higher order derivatives in the same way:

$$\frac{\partial^k f}{\partial x_1\,\partial x_2\,\ldots\,\partial x_k}.$$

Example: if $f(x_1, x_2) = 2x_1^2x_2^4 - 5x_1x_2^3 + 16$ , then

$$f''_{x_1,x_1} = 4x_2^4, \quad f''_{x_1,x_2} = 16x_1x_2^3 - 15x_2^2, \quad f''_{x_2,x_2} = 24x_1^2x_2^2 - 30x_1x_2, \quad f'''_{x_1,x_1,x_2} = 16x_2^3, \ldots$$