Directional derivative and gradient vector

- Definition of directional derivative. (Sec. 14.6)
- Directional derivative and partial derivatives.
- Gradient vector.
- Geometrical meaning of the gradient.

The directional derivative generalizes the partial derivatives to any direction

**Definition 1** The directional derivative of the function $f(x, y)$ at the point $(x_0, y_0)$ in the direction of a unit vector $\mathbf{u} = (u_x, u_y)$ if

$$D_{\mathbf{u}}f(x_0, y_0) = \lim_{t \to 0} \frac{1}{t} \left[ f(x_0 + u_xt, y_0 + u_yt) - f(x_0, y_0) \right],$$

if the limit exists.

Particular cases:
- $\mathbf{u} = \langle 1, 0 \rangle = \mathbf{i}$, then $D_{\mathbf{i}}f(x_0, y_0) = f_x(x_0, y_0)$.
- $\mathbf{u} = \langle 0, 1 \rangle = \mathbf{j}$, then $D_{\mathbf{j}}f(x_0, y_0) = f_y(x_0, y_0)$. 
Recall the definition of partial derivatives

The directional derivative generalizes the partial derivatives to any direction
\[ |u| = 1 \text{ implies that } t \text{ is the distance between the points } (x, y) = (x_0 + u_x t, y_0 + u_y t) \text{ and } (x_0, y_0) \]

\begin{align*}
    d &= |\langle x - x_0, y - y_0 \rangle|, \\
    &= |\langle u_x t, u_y t \rangle|, \\
    &= |t| |u|, \\
    &= |t|.
\end{align*}

The directional derivative of \( f(x, y) \) at \((x_0, y_0)\) along \( u \) is the pointwise rate of change of \( f \) with respect to the distance along the line parallel to \( u \) passing through \((x_0, y_0)\).

---

Here is a useful formula to compute directional derivative

**Theorem 1** If \( f(x, y) \) is differentiable and \( u = \langle u_x, u_y \rangle \) is a unit vector, then

\[ D_u f(x_0, y_0) = f_x(x_0, y_0) u_x + f_y(x_0, y_0) u_y. \]

The proof is based in the chain rule, case 1
Proof of the theorem

Chain rule case 1, for \( x(t) = x_0 + u_x t, \ y(t) = y_0 + u_y t \).
Then, \( z(t) = f(x(t), y(t)) \).

On the one hand,
\[
\left. \frac{dz}{dt} \right|_{t=0} = \lim_{t \to 0} \frac{1}{t} [z(t) - z(0)],
\]
\[
= \lim_{t \to 0} \frac{1}{t} [f(x_0 + u_x t, y_0 + u_y t) - f(x_0, y_0)],
\]
\[
= D_u f(x_0, y_0).
\]

Proof of the theorem (Cont.)

On the other hand,
\[
\frac{dz}{dt}(t) = f_x(x(t), y(t)) \frac{dx}{dt}(t) + f_y(x(t), y(t)) \frac{dy}{dt}(t),
\]
then,
\[
\left. \frac{dz}{dt} \right|_{t=0} = f_x(x_0, y_0) u_x + f_y(x_0, y_0) u_y.
\]
Therefore,
\[
D_u f(x_0, y_0) = f_x(x_0, y_0) u_x + f_y(x_0, y_0) u_y.
\]
Example about how to compute a directional derivative

Let \( f(x, y) = \sin(x + 2y) \). Compute the directional derivative of \( f(x, y) \) at \((4, -2)\) in the direction \( \theta = \pi/6 \).

\[
\mathbf{u} = (\cos(\theta), \sin(\theta)), \quad \mathbf{u} = (\sqrt{3}/2, 1/2).
\]

Also
\[
f_x = \cos(x + 2y), \quad f_y = 2\cos(x + 2y),
\]
then
\[
D_u f(x, y) = \cos(x + 2y)u_x + 2\cos(x + 2y)u_y,
\]
\[
D_u f(4, -2) = \frac{\sqrt{3}}{2} + 1.
\]

Directional derivatives can be defined on functions of 2, 3 or more variables

**Definition 2** (functions of 3 variables)

*The directional derivative of the function \( f(x, y, z) \) at the point \((x_0, y_0, z_0)\) in the direction of a unit vector \( \mathbf{u} = (u_x, u_y, u_z) \) is*

\[
D_u f(x_0, y_0, z_0) = \lim_{t \to 0} \frac{1}{t} \left[ f(x_0 + u_x t, y_0 + u_y t, z_0 + u_z t) - f(x_0, y_0, z_0) \right],
\]

*if the limit exists.*
The same useful theorem we had in 2 variable functions

**Theorem 2** If \( f(x, y, z) \) is differentiable and 
\( u = \langle u_x, u_y, u_z \rangle \) is a unit vector, then
\[
D_u f(x_0, y_0, z_0) = f_x(x_0, y_0, z_0) u_x + f_y(x_0, y_0, z_0) u_y + f_z(x_0, y_0, z_0) u_z.
\]

The directional derivative can be written in terms of a dot product

In the case of 2 variable functions:
\[
D_u f = f_x u_x + f_y u_y = (\nabla f) \cdot u,
\]
with \( \nabla f = \langle f_x, f_y \rangle \).

In the case of 3 variable functions:
\[
D_u f = f_x u_x + f_y u_y + f_z u_z = (\nabla f) \cdot u,
\]
with \( \nabla f = \langle f_x, f_y, f_z \rangle \).
We introduce the gradient vector for functions of 2 or 3 variables

**Definition 3** Let \( f(x, y, z) \) be a differentiable function. Then,

\[
\nabla f(x, y, z) = (f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)),
\]

is called the gradient of \( f(x, y, z) \).

In 2 variables: \( \nabla f(x, y) = (f_x(x, y), f_y(x, y)) \).

Alternative notation: \( \nabla f = f_x \mathbf{i} + f_y \mathbf{j} + f_z \mathbf{k} \).

The useful theorem now has the following form

**Theorem 3** Let \( f(x, y, z) \) be differentiable function. Then,

\[
D_u f(x) = (\nabla f(x)) \cdot u.
\]

with \(|u| = 1\).
Gradient vector  The gradient vector has two main properties:

- It points in the direction of the maximum increase of \( f \), and \(|\nabla f|\) is the value of the maximum increase rate.
- \( \nabla f \) is normal to the level surfaces.

Here is the first property of the gradient vector

**Theorem 4** Let \( f \) be a differentiable function of 2 or 3 variables. Fix \( P_0 \in D(f) \), and let \( u \) be an arbitrary unit vector.

Then, the maximum value of \( D_u f(P_0) \) among all possible directions is \(|\nabla f(P_0)|\), and it is achieved for \( u \) parallel to \( \nabla f(P_0) \).
The proof of the first property

\[ D_u f(P_0) = (\nabla f(P_0)) \cdot u, \]
\[ = |\nabla f(P_0)| |u| \cos(\theta), \]
\[ = |\nabla f(P_0)| \cos(\theta). \]

But \(-1 \leq \cos(\theta) \leq 1\) implies
\[ -|\nabla f(P_0)| \leq D_u f(P_0) \leq |\nabla f(P_0)|. \]

And \(D_u f(P_0) = |\nabla f(P_0)|, \Leftrightarrow \theta = 0 \Leftrightarrow u\) is parallel \(\nabla f(P_0)\).

Here is the second property of the gradient vector, in the case of 3 variable functions

**Theorem 5** Let \(f(x, y, z)\) be a differentiable at \(P_0\). Then, \(\nabla f(P_0)\) is orthogonal to the plane tangent to a level surface containing \(P_0\).
Proof of the second property

Let \( \mathbf{r}(t) \) be any differentiable curve in the level surface \( f(x, y, z) = k \). Assume that \( \mathbf{r}(t = 0) = \overrightarrow{OP}_0 \). Then,

\[
0 = \frac{df}{dt}, \\
= f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}, \\
= \nabla f(\mathbf{r}(t)) \cdot \frac{d\mathbf{r}}{dt}(t).
\]

But \( (d\mathbf{r})/(dt) \) is tangent to the level surface for any choice of \( \mathbf{r}(t) \).

Therefore

\[
\nabla f(\mathbf{r}(t = 0)) \cdot \frac{\mathbf{r}}{dt}(t = 0) = 0
\]

implies that \( \nabla f(P_0) \) is orthogonal to the level surface.

---

Local and absolute maxima, minima, and inflection points

- Definitions of local extrema. (Sec. 14.7)
- Characterization of local extrema.
- Absolute extrema on closed and bounded sets.
- Typical exercises.
Recall the main results on local extrema for $f(x)$

- At $f$, $f'$, $f''$
  - $a$ max. $0 < 0$
  - $b$ infl. $\neq 0 \pm 0\mp$
  - $c$ min. $0 > 0$
  - $d$ infl. $= 0 \pm 0\mp$

The main cases of local extrema for $f(x, y)$
The intuitive notions of local extrema can be written precisely as follows

**Definition 4 (Local maximum)** A function \( f(x, y) \) has a local maximum at \((a, b) \in D(f) \) if \( f(x, y) \leq f(a, b) \) for all \((x, y)\) near \((a, b)\).

**Definition 5 (Local minimum)** A function \( f(x, y) \) has a local minimum at \((a, b) \in D(f) \) if \( f(x, y) \geq f(a, b) \) for all \((x, y)\) near \((a, b)\).

The tangent plane to the graph of \( f \) at a local max-min is horizontal

**Theorem 6** Let \( f(x, y) \) be differentiable at \((a, b)\). If \( f \) has a local maximum or minimum at \((a, b)\) then \( \nabla f(a, b) = (0, 0) \).

Recall: \( n = \langle f_x, f_y, -1 \rangle = \langle 0, 0, -1 \rangle \).

The converse is not true: It could be a saddle point
Stationary points include local maxima, minima, and saddle points

**Definition 6 (Stationary point)** Let \( f(x, y) \) be a differentiable function at \((a, b)\). If \( \nabla f(a, b) = (0, 0) \), then the point \((a, b)\) is called a stationary point of \( f \).

Stationary point are located where the gradient vector vanishes

**Theorem 7 (Second derivative test)** Let \((a, b)\) be a stationary point of \( f(x, y) \), that is, \( \nabla f(a, b) = 0 \). Assume that \( f(x, y) \) has continuous second derivatives in a disk with center in \((a, b)\). Introduce the quantity

\[
D = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2.
\]

- If \( D > 0 \) and \( f_{xx}(a, b) > 0 \), then \( f(a, b) \) is a local minimum.
- If \( D > 0 \) and \( f_{xx}(a, b) < 0 \), then \( f(a, b) \) is a local maximum.
- If \( D < 0 \), then \( f(a, b) \) is a saddle point.
- If \( D = 0 \) the test is inconclusive.
Find the local extrema of \( f(x, y) = y^2 - x^2 \)

\[ \nabla f = (-2x, 2y), \quad \Rightarrow \quad \nabla f = (0, 0) \quad \text{at} \quad (0, 0). \]

\[ f_{xx}(0, 0) = -2, \quad f_{yy}(0, 0) = 2, \quad f_{xy}(0, 0) = 0; \]

\[ D = (-2)(2) = -4 < 0 \quad \Rightarrow \quad \text{saddle point at} \quad (0, 0). \]

Is \((0,0)\) a local extrema of \( f(x, y) = y^2x^2 \)?

\[ \nabla f(x, y) = (2xy^2, 2yx^2), \quad \Rightarrow \quad \nabla f(0, 0) = (0, 0) \quad \text{at} \quad (0, 0). \]

\[ f_{xx}(x, y) = 2y^2, \quad f_{yy}(x, y) = 2x^2, \quad f_{xy}(x, y) = 4xy, \]

\[ f_{xx}(0, 0) = 0, \quad f_{yy}(0, 0) = 0, \quad f_{xy}(0, 0) = 0, \]

So \( D = 0 \) and the test is inconclusive.
From the graph of $f = x^2y^2$ is easy to see that $(0, 0)$ is a global minimum.

Find the maximum volume of a closed rectangular box with a given surface area $A_0$

$$V(x, y, z) = xyz, \quad A(x, y, z) = 2xy + 2xz + 2yz.$$  

But $A(x, y, z) = A_0$, then

$$z = \frac{A_0 - 2xy}{2(x + y)}, \quad \Rightarrow \quad V(x, y) = \frac{A_0xy - 2x^2y^2}{2(x + y)}.$$  

Find $\nabla V(x_0, y_0) = (0, 0)$.  

The result is $x_0 = y_0 = z_0 = \sqrt{A_0/6}$.  

Local extrema need not be the absolute extrema

Absolute extrema may not be defined on open intervals

Continuous functions $f(x)$ on intervals $[a, b]$ always have absolute extrema

Intervals $[a, b]$ are bounded and closed sets in $\mathbb{R}$

Because they do not extend to infinity, and the boundary points belong to the set.
Here is the generalization of closed and bounded intervals to $\mathbb{R}^2$

**Definition 7** A set $D \subset \mathbb{R}^2$ is bounded if it can be contained in a disk. A set $D \in \mathbb{R}^2$ is closed if it contains all its boundary points.

A point $P \in \mathbb{R}^2$ is a boundary point of a set $D$ if every disk with center in $P$ always contains both points in $D$ and points not in $D$.

Here are examples of bounded sets

- $\{x^2 + y^2 \leq 1\}$, Closed and bounded.
- $\{x^2 + y^2 < 1\}$, Open and bounded.
Here are more examples of different types of sets

- Open and bounded
- Closed and bounded
- Bounded
- Closed and unbounded

Continuous functions on bounded and closed sets always have absolute extrema

**Theorem 8** If $f(x, y)$ is continuous in a closed and bounded set $D \subset \mathbb{R}^2$, then $f$ has an absolute maximum and an absolute minimum in $D$. 
Suggestions to find absolute extrema of \( f(x, y) \) in a closed and bounded set

- Find every stationary point of \( f \).
  \((\nabla f(x, y) = 0). \) No second derivative test needed.)
- Find the extrema (max. and min.) values of \( f \) on the boundary of \( D \).
- The biggest (smallest) of the previous steps is the absolute maximum (minimum).

Here is a typical exercise

Find the absolute extrema of \( f(x, y) = 4x + 6y - x^2 - y^2 \), on \( D = \{(x, y) \in \mathbb{R}^2, \; 0 \leq x \leq 4, \; 0 \leq y \leq 5\} \)

Absolute minimum: \((4, 0), (0, 0)\).
Absolute maximum: \((2, 3)\).
Lagrange’s multipliers

- Example of the method.
- Maximization of functions subject to constraints.
- Examples.
- Generalization to more than one constraint.

Example: Find the rectangle of biggest area with fixed perimeter $P_0$

One way to solve the problem is:

$$A(x,y) = xy, \quad P_0 = P(x, y) = 2x + 2y,$$

then $y = P_0/2 - x$, and replace it in $A(x, y)$,

$$A(x) = \frac{P_0}{2}x - x^2.$$

The stationary points of this function are

$$0 = A'(x) = \frac{P_0}{2} - 2x, \quad \Rightarrow \quad x = \frac{P_0}{4}, \quad \Rightarrow \quad y = \frac{P_0}{4}.$$

So the answer is the square of side

$$x = y = \frac{P_0}{4}.$$
Idea behind the Lagrange multipliers method

Level curves of $A = xy$, Level curves of the constraint $P = 2x + 2y$.

The gradient vectors of $A(x, y)$ and of the constraint $P = 2x + 2y$ are parallel at the solution.
The same problem solved with the Lagrange multipliers method

Find the maximum of \( A(x, y) = xy \) subject to the constraint \( P(x, y) = 2x + 2y = P_0 \).

One has to find the \((x, y)\) such that

\[
\nabla A(x, y) = \lambda \nabla P(x, y), \quad P(x, y) = P_0,
\]

with \( \lambda \neq 0 \). From the first equation one has

\[
(y, x) = \lambda (2, 2), \quad \Rightarrow \quad x = 2\lambda, y = 2\lambda.
\]

Then the constraint \( P_0 = 2x + 2y \) implies that \( P_0 = 8\lambda \), so the answer is

\[
x = y = \frac{P_0}{4}.
\]

Lagrange multipliers method can be summarized as follows:

The extrema values of \( f(x, y) \) subject to the constraint \( g(x, y) = k \) can be obtained as follows:

- Find all solutions \((x_0, y_0)\) and \( \lambda \) of the equations

\[
\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0),
\]

\[
g(x_0, y_0) = k.
\]

- Evaluate \( f \) at every solution \((x_0, y_0)\). The largest and smallest values are respectively the maximum and minimum values of \( f \) subject to the constraint \( g = k \).
Lagrange multipliers method for functions of three variables

The extrema values of \( f(x, y, z) \) subject to the constraint \( g(x, y, z) = k \) can be obtained as follows:

- Find all solutions \((x_0, y_0, z_0)\) and \( \lambda \) of the equations
  \[
  \nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0),
  \]
  \[
  g(x_0, y_0, z_0) = k.
  \]

- Evaluate \( f \) at every solution \((x_0, y_0, z_0)\). The largest and smallest values are respectively the maximum and minimum values of \( f \) subject to the constraint \( g = k \).

Example: Find the rectangular box of maximum volume for fixed area.

The function is \( V(x, y, z) = xyz \). The constraint function is \( A(x, y, z) = 2xy + 2xz + 2yz \). The constraint is \( A(x, y, z) = A_0 \).

Find the \((x, y, z)\) solutions of
\[
\nabla V = \lambda \nabla A,
\]
\[
A = A_0.
\]

These equations are:
\[
yz = 2\lambda(z + y),
xz = 2\lambda(x + z),
xy = 2\lambda(x + y),
A_0 = 2(xy + xz + yz).
\]

The solution is \( x = y = z = \sqrt{A_0/6} \).
Example: Find the extrema values of
\[ f(x, y) = x^2 + y^2/4 \] in the circle \( x^2 + y^2 = 1 \)

Then, \( f(x, y) = x^2 + y^2/4 \), and \( g(x, y) = x^2 + y^2 \). The equations are:
\[
\begin{align*}
\nabla f &= \lambda \nabla g, \\
\Rightarrow \quad 2x &= \lambda (2x/2), \\
g &= 1, \\
\Rightarrow \quad x^2 + y^2 &= 1.
\end{align*}
\]
Which imply
\[
\begin{align*}
x &= \lambda x, \\
\Rightarrow \quad (1 - \lambda)x &= 0, \\
y/2 &= 2\lambda y, \\
\Rightarrow \quad (1/4 - \lambda)y &= 0, \\
x^2 + y^2 &= 1.
\end{align*}
\]
The solutions are: \( P = (0, \pm 1) \), and \( P = (\pm 1, 0) \). Then:
\[
f(0, \pm 1) = 1/4, \text{ absolute minimum in the circle.}
\]
\[
f(\pm 1, 0) = 1, \text{ absolute maximum in the circle.}
\]

Generalization to two constraints
The extrema values of \( f(x, y, z) \) subject to the constraints \( g(x, y, z) = k_1 \) and \( h(x, y, z) = k_2 \) can be obtained as follows:

- Find all solutions \((x_0, y_0, z_0)\) and \(\lambda\) of the equations
  \[
  \nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0) + \mu \nabla h(x_0, y_0, z_0), \\
g(x_0, y_0, z_0) = k_1, \\
h(x_0, y_0, z_0) = k_2.
  \]
- Evaluate \( f \) at every solution \((x_0, y_0, z_0)\). The largest and smallest values are respectively the maximum and minimum values of \( f \) subject to the constraint \( g = k_1 \) and \( h = k_2 \).