Lagrange multipliers:

Problem: Find extrema of $f(x)$ under some constraint $g(x) = 0$

Approach via differentials:

- for $f$ to have extremum we need $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0$
- but $dx, dy$ not independent, since they are such that $0 = \partial g = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy$

$\Rightarrow$ $0 = df + \lambda dg = (\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x}) dx + (\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y}) dy$ for any $\lambda \in \mathbb{R}$

regarded as independent now, since constraint taken into account

$\Rightarrow$ need to solve $\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0$ and $\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0$

More geometric approach:

- consider $f$ restricted to $S = \{ x \in \mathbb{R}^n : g(x) = 0 \}$ with an extremum at $\bar{a}$

and any curve $\tilde{h}(t)$ in $S$ such that $\tilde{h}(0) = \bar{a}$.

$\Rightarrow$ $v(t) = f(\tilde{h}(t))$ has an extremum at $t = 0$

$\Rightarrow$ $0 = v'(0) = (\tilde{\nabla} f)(\bar{a}) \cdot \tilde{h}'(0)$

$\Rightarrow (\tilde{\nabla} f)(\bar{a})$ orthogonal to tangent vector of any curve in $S$

but $(\tilde{\nabla} g)(\bar{a})$ is also orthogonal to $S$

$\Rightarrow (\tilde{\nabla} f)(\bar{a}) = \lambda (\tilde{\nabla} g)(\bar{a})$ for some $\lambda \in \mathbb{R}$ (both vectors are linearly dependent)
**Conclusion (Lagrange's method):**

As necessary condition to find extrema of \( f(x) \) under constraint \( g(x) = 0 \), we need to solve

\[
\nabla (f - \lambda g) = \mathbf{0} \quad \text{and} \quad g(x) = 0 \quad \text{i.e.,} \ n+1 \text{ equations for } n+1 \text{ variables } x_1, \ldots, x_n, \lambda.
\]

\( \lambda \) is called Lagrange multiplier.

In short: if \( L(x, \lambda) = f(x) - \lambda g(x) \), we need \( \nabla L = 0 \)

\[
\begin{pmatrix}
\frac{\partial L}{\partial x_1} \\
\frac{\partial L}{\partial x_2} \\
\vdots \\
\frac{\partial L}{\partial \lambda}
\end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}
\]

**Ex.: rectangle:** \( f(x, y) = xy \), \( g(x, y) = 2x + 2y - L = 0 \)

\[
\Rightarrow L(x, y, \lambda) = xy - \lambda (2x + 2y - L)
\]

\[
\Rightarrow \nabla L = \begin{pmatrix}
\frac{\partial L}{\partial x} \\
\frac{\partial L}{\partial y} \\
\frac{\partial L}{\partial \lambda}
\end{pmatrix} = \begin{pmatrix} y - 2\lambda \\ x - 2\lambda \\ 2x + 2y - L \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\]

\[
\Rightarrow 0 = 2x + 2y - L = 2(2\lambda) + 2(2\lambda) - L = 0 \quad \Rightarrow \lambda = \frac{L}{8} \quad \Rightarrow x = \frac{L}{4}, \quad y = \frac{L}{4}
\]

usually not interested in the value of \( \lambda \)

\( f(x, y) = x^2 + y^2 + 1 \), \( g(x, y) = x^2 + y^2 - 1 = 0 \) **(circle)**

\[
\Rightarrow L = x^2 y + y - \lambda (x^2 + y^2 - 1) \Rightarrow \begin{pmatrix}
\frac{\partial L}{\partial x} \\
\frac{\partial L}{\partial y} \\
\frac{\partial L}{\partial \lambda}
\end{pmatrix} = \begin{pmatrix} 2x - 2\lambda x \\ 2y + 2\lambda y \\ x^2 + y^2 - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\]
\[ x = 0 \text{ or } \lambda = 1 \implies \text{only } x = 0 \implies \gamma = \pm \sqrt{x^2 - 1} = \pm 1 \]
\[ \implies 2\gamma + 1 - 2\gamma = 1 \text{ cannot be zero} \]

\[ \implies (0, -1) \text{ and } (0, 1) \text{ are critical points} \]
\[ \downarrow \quad \downarrow \]
\[ f(0, -1) = 0 \quad , \quad f(0, 1) = 2 \]
\[ \implies \text{indeed a minimum} \implies \text{indeed a maximum} \]

**Note:** There are at least two advantages of using Lagrange's method:

- We can use it even if we cannot solve the constraint for one of the variables, or generally when substituting it into the function does not work.

- Even if we could substitute the constraint into the function, and then find the critical points as usual, the computation with Lagrange's method is usually easier and faster, see examples in the homework/moodle/quiz exercises.
2.5 Vector-valued Functions

We know what derivatives of \( f: \mathbb{R} \to \mathbb{R} \) and \( f: \mathbb{R}^n \to \mathbb{R} \) are.

What about \( f: \mathbb{R} \to \mathbb{R}^n \)?

\[
\vec{f}(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix} \implies \frac{d\vec{f}}{dt} := \lim_{h \to 0} \frac{\vec{f}(t+h) - \vec{f}(t)}{h} = \lim_{h \to 0} \begin{pmatrix} \frac{1}{h} (f_1(t+h) - f_1(t)) \\ \vdots \\ \frac{1}{h} (f_n(t+h) - f_n(t)) \end{pmatrix},
\]

so \( \frac{d\vec{f}}{dt} = \begin{pmatrix} \frac{df_1}{dt} \\ \vdots \\ \frac{df_n}{dt} \end{pmatrix} \) i.e., derivative is taken component-wise.

\( \Rightarrow \) the usual product rules hold: for \( \vec{f}, \vec{g}: \mathbb{R} \to \mathbb{R}^n \), \( \Phi: \mathbb{R} \to \mathbb{R} \), then

\[
\begin{align*}
\frac{d}{dt} (\Phi \vec{f}) &= \frac{d\Phi}{dt} \vec{f} + \Phi \frac{d\vec{f}}{dt} \\
\frac{d}{dt} (\vec{f} \cdot \vec{g}) &= \frac{d\vec{f}}{dt} \cdot \vec{g} + \vec{f} \cdot \frac{d\vec{g}}{dt} \\
\frac{d}{dt} (\vec{f} \times \vec{g}) &= \frac{d\vec{f}}{dt} \times \vec{g} + \vec{f} \times \frac{d\vec{g}}{dt}
\end{align*}
\]

Recall: cross product \( \vec{f} \times \vec{g} = \begin{pmatrix} f_2 g_3 - f_3 g_2 \\ f_3 g_1 - f_1 g_3 \\ f_1 g_2 - f_2 g_1 \end{pmatrix} \)