Overview no. 1

The topics for the course on optimisation will this autumn be:

- Extrema under constraints for real-valued functions.
- Calculus of Variations.
- Optimal Control Theory.

The purpose of the course is to acquaint you with the topics, their methods and the kind *problems* they can solve (or address).

As you will see, the problems go far beyond what you could treat previously by setting gradients equal to zero. Phrased briefly, we shall deal with general, yet powerful methods for rewriting optimisation problems in terms of ordinary differential equations, which can be solved. But more about this later.

The course will be based on the book (ask secr. Lisbeth G. Nielsen)

[SS] *Optimal control theory with economic applications,* by Atle Seierstad og Knut Sydsæter; North Holland 1987.

It will be a central issue in this course to work through the exercises. However, it will probably be best to delay the exercises to the next gathering after the lectures on a subject. I therefore propose the following:

1st gathering, Thursday September 2. We meet at 12.30-14.15 in G5-109 for the lectures. I will begin with an overview of the course, including a primer on Calculus of Variations and on Optimal Control theory. Then we go through the first topic, which is a deeper understanding of extrema of real-valued functions. This includes necessary conditions for extrema (second order derivatives) and extrema under *constraints* — think for example of finding the maximum of a function defined on the unit sphere in $\mathbb{R}^n$.

From 14.30 to 16.15 you may begin in the groups to solve the exercises announced below for the second gathering.

For convenience the next overview contains a few notes for today's lectures.

Med venlig hilsen
Jon Johnsen

Overview no. 2

For completeness we first give a precise version of Taylor's formula: when $g\colon I \to \mathbb{C}$ is a C^n-function on an open interval $I \subset \mathbb{R}$, then it holds for all t, $t_0 \in I$ that

$$g(t) = g(t_0)+\dots+\tfrac{g^{(n-1)}(t_0)}{(n-1)!}(t-t_0)^{n-1}+\tfrac{(t-t_0)^n}{(n-1)!}\int_0^1 (1-\theta)^{n-1}g^{(n)}(t_0+\theta(t-t_0))\,d\theta. \tag{1}$$

This can be verified by induction over n, by using that $(1-\theta)^{n-1} = -\frac{1}{n}\frac{d}{d\theta}(1-\theta)^n$.

The above gives easy access to *Taylor's limit formula* for $g \in C^n(I)$:

$$g(t) = g(t_0) + \dots + \tfrac{g^{(n)}(t_0)}{n!}(t-t_0)^n + o((t-t_0)^n), \tag{2}$$

where by definition $o(s^n)/s^n \to 0$ for $s \to 0$. In fact, the o-function can be obtained from (1) by subtraction of $\frac{g^{(n)}(t_0))}{n!}(t-t_0)^n$ from the integral remainder, for $\frac{1}{n} = \int_0^1 (1-\theta)^{n-1}\,d\theta$ can be used to exploit the continuity of $g^{(n)}$ at t_0 (try it !).

For functions f of several variables, both formulae can be utilised eg by introducing $g(t) = f(x^* + t(x-x^*))$, when x is in a ball $B(x^*, r)$.

Theorem. When $f\colon O \to \mathbb{R}$ be C^2 on an open set $O \subset \mathbb{R}^n$, then

$$f \text{ has a local maximum at } x^* \in O \implies \begin{cases} \nabla f(x^*) = 0, \\ \lambda \le 0 \text{ for each eigenvalue } \lambda \text{ for} \\ Hf(x^*) := (\frac{\partial^2 f}{\partial x_j \partial x_k}(x^*)). \end{cases} \tag{3}$$

Conversely, if $\lambda < 0$ whenever λ is an eigenvalue for $Hf(x^*)$, whereby $x^* \in O$ is a critical point, then x^* is a local maximum.

Note that there is a small gap between the necessary and sufficient conditions for the local maximum; this cannot be avoided, cf the exercise below.

Proof: Since $f \in C^2$, Taylor's formula yields that

$$f(x)-f(x^*) = \nabla f(x^*)\cdot(x-x^*)+\tfrac{1}{2}(x-x^*)^T Hf(x^*)(x-x^*)+o(\|x-x^*\|^2). \tag{4}$$

When x^* is a local extremum, this gives $\nabla f(x^*) = 0$. For otherwise one can take $\|x-x^*\|$ outside a parenthesis on the right-hand side and note that

$$x \mapsto \nabla f(x^*) \cdot (\tfrac{1}{\|x-x^*\|}(x-x^*))$$

takes on both positive and negative values on every line segment parallel to $\nabla f(x^*)$ through x^*; on suitably small such segments this also holds after addition of (two) terms of the form $o(\|x-x^*\|)$. This contradicts that $f(x) - f(x^*)$ has constant sign on such line segments. Thence $\nabla f(x^*) = 0$.

The matrix Hf is real and symmetric, as $f \in C^2$. So according to the spectral theorem, $Hf(x^*)$ is diagonalisable and has real eigenvalues $\lambda_1, \dots, \lambda_n$. Thus $Hf(x^*) = PDP^T$ for $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ and a suitably chosen $n \times n$-matrix P containing an orthonormal basis of eigenvectors for $Hf(x^*)$. Setting $y = P^T(x - x^*)$, whereby $\|y\| = \|x - x^*\|$, this gives

$$\begin{aligned} f(x) - f(x^*) &= \tfrac{1}{2}(x - x^*)^T PDP^T (x - x^*) + o(\|x - x^*\|^2) \\ &= \tfrac{1}{2}(\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2) + o(\|y\|^2). \end{aligned} \tag{5}$$

In particular $y = (0, \dots, 0, y_j, 0, \dots, 0)$ with corresponding $x - x^* = Py$, entails

$$0 \geq f(x) - f(x^*) = (\tfrac{1}{2}\lambda_j + o(1)) y_j^2 \tag{6}$$

for $\|x - x^*\| = |y_j|$ in a neighbourhood of 0, since x^* is a maximum point. As $\frac{1}{2}\lambda_j + o(1)$ has the same sign as λ_j for $|y_j|$ in a possibly smaller neighbourhood of 0, it follows that $\lambda_j > 0$ is impossible. Hence $\lambda_j \leq 0$ for $j = 1, \dots, n$.

Conversely, note that (5) implies

$$f(x) - f(x^*) \leq (\tfrac{1}{2} \max(\lambda_1, \dots, \lambda_n) + o(1)) \|y\|^2. \tag{7}$$

If $\lambda_j < 0$ for all j, the parenthesis is negative or 0 for y in a small ball $B(0, r)$. This entails that $f(x) \leq f(x^*)$ for $x \in B(x^*, r)$, as desired.

To elucidate the theorem, it is recalled that a quadratic form Q on $\mathbb{R}^n$ is a function of the form

$$Q(x) = x^T A x,$$

whereby A is a fixed $n \times n$-matrix. In practice A is often real and symmetric.

By definition, Q is said to be

$$\text{positive definite if } x^T A x > 0 \text{ for every } x \neq 0; \tag{8}$$

$$\text{positive semidefinite if } x^T A x \geq 0 \text{ for every } x \in \mathbb{R}^n. \tag{9}$$

Similarly Q is negative definite and negative semidefinite, if $-Q$ is positive definite and positive semidefinite, respectively. Moreover, Q is called

$$\text{indefinite if for some } x, y \in \mathbb{R}^n \text{ it holds that } Q(x) > 0 > Q(y). \tag{10}$$

The same terminology applies to the matrix A defining the quadratic form Q.

These notions can be effectively analysed in terms of eigenvalues. In fact, when A is reall and symmetric, a diagonalisation of $x^t A x$, similar to the one in (5), shows straightforwardly that

$$Q \text{ is positive definite } \iff \lambda > 0 \text{ holds for every eigenvalue } \lambda \text{ of } A \tag{11}$$

$$Q \text{ is positive semidefinite } \iff \lambda \geq 0 \text{ holds for every eigenvalue } \lambda \text{ of } A. \tag{12}$$

Of these inequalities are reversed for the corresponding versions of negative definiteness. Moreover,

$$Q \text{ is indefinite } \iff \mu < 0 < \lambda \text{ holds for two eigenvalues } \mu, \lambda \text{ of } A. \tag{13}$$

The *signature* of Q, or of A, is then given by

$$\operatorname{sign}(A) = \dim V_+ - \dim V_-. \tag{14}$$

Here V_+ is the maximal subspace on which (the restriction of) A has only positive eigenvalues; similarly for V_-. Clearly $\mathbb{R}^n = V_+ \oplus \operatorname{Null} A \oplus V_-$.

Returning to the theorem, it is clear that for a critical oint $x^* \in O$ to be a local maximum, it is necessary that the Hessian $Hf(x^*)$ is negative *semidefinite*. The converse is not assured; but it is sufficient for a maximum that $Hf(x^*)$ is negative definite.

Remark. The definition of a *saddle point* is very concise in these terms: it is a critical point at which the Hessian is indefinite. So if x^* is a saddle point, then $\nabla f(x^*) = 0$ and $Hf(x^*)$ has at least one positive and one negative eigenvalue — therefore there are two lines through x^* (along the corresponding eigenvectors) on which f has a minimum at x^*, respectively a maximum at x^*. (Visualise a horse saddle!)

Remark. At a critical point $x^* \in O$, there are consequently four possibilities:

- f has a local maximum at x^*;
- f has a local minimum at x^*;
- f has a saddle point at x^*;
- f has a more complicated behaviour in every neighbourhood of x^*.

The last point should not be overlooked. One simple example of this is provided by the fourth order polynomial on $\mathbb{R}^2$, $f(x, y) = (y - x^2)(y - 3x^2)$: along the first axis this has a minimum at $(0, 0)$; yet along any other line through the origin it attains both positive and negative values, even arbitrarily close to $(0, 0)$; hence f has neither an extremum at $(0, 0)$ nor a saddle point there.

As a reminder from mathematics 2 we also have

The Implicit Function theorem. When $\Phi(x, y)$ is a C^1-function $O \to \mathbb{R}^k$, on an open set $O \subset \mathbb{R}^{n-k} \times \mathbb{R}^k$, and (x_0, y_0) solves the equation

$$\Phi(x, y) = 0,$$

and moreover the Jacobian $D_y\Phi$ of $\Phi(x_0, \cdot)$ at the point y_0 has an inverse $\frac{\partial \Phi}{\partial y}(x_0, y_0)^{-1}$, then there are closed balls $U = \bar{B}(x_0, \alpha)$ and $V = \bar{B}(y_0, \beta)$ and a function $\psi \in C^1(U, V)$ so that all the equation's solutions (x, y) in $U \times V$ constitute the graph of ψ. In other words,

$$\{\, (x, \psi(x)) \mid x \in U \,\} = (U \times V) \bigcap \Phi^{-1}(\{0\}). \tag{15}$$

Since $\Phi(x, \psi(x)) \equiv 0$ the chain rule gives, in block matrix notation,

$$\begin{pmatrix} D_x\Phi & D_y\Phi \end{pmatrix} \begin{pmatrix} I \\ D\psi \end{pmatrix} = 0. \tag{16}$$

This theorem is a valuable tool in the next section.

Optimisaton under constraints. In the following $f\colon O \to \mathbb{R}$ denotes a C^1-function on an open set $O \subset \mathbb{R}^n$. The task will be to determine its extreme values as x runs through a subset $\mathcal{F} \subset O$, given as the set of solutions (in O, of course) to the equations

$$g_1(x) = c_1, \quad \dots \quad , g_k(x) = c_k. \tag{17}$$

Hereby the $g_j \in C^1(O, \mathbb{R})$ are given functions. To determine $\max_{x\in\mathcal{F}} f(x)$ and $\min_{x\in\mathcal{F}} f(x)$ is called maximisation and minimisation, respectively, of f under the constraints (17) given by $g_1,\dots,g_k$.

Henceforth it is assumed that $k < n$, for otherwise one can only expect to have (none or) finitely many solutions to (17).

Notice that the Jacobian of $\mathcal{G} := \begin{pmatrix} g_1 \\ \vdots \\ g_k \end{pmatrix}$ is not assumed surjective, so the pre-image $\mathcal{F} = \mathcal{G}^{-1}(\{(c_1, \dots, c_k)\})$ is therefore not necessarily a regular C^1-surface in O. (Cf the mathematics 3 course in geometry.)

As a necessary condition for extremum under constraints one has:

Theorem. If $x^* \in O$ satisfies the constraints (17) and gives f an extreme value on $\mathcal{F}$, then the matrix

$$M = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial x_1} & \dots & \frac{\partial g_k}{\partial x_n} \end{pmatrix} (x^*) \tag{18}$$

does *NOT* have maximal rank. (Ie there is at most k linearly independent rows.)

Proof: Suppose $\operatorname{rank} M = k + 1$; by relabelling the variables it may be assumed that the first $k + 1$ columns are linearly independent.

Introducing an auxiliary variable $u \in \mathbb{R}$, it is convenient to set

$$F(x, u) = \begin{pmatrix} f(x) + u \\ g_1(x) \\ \vdots \\ g_k(x) \end{pmatrix}.$$

This is a C^1-function $O \times \mathbb{R} \to \mathbb{R}^{k+1}$ with $\frac{\partial F}{\partial x, \partial u}(x^*, 0) = \begin{pmatrix} M & e_1 \end{pmatrix}$, in block matrix notation. Hereby e_1 denotes the first canonical basis vector.

To apply the Implicit Function Theorem, the points $(x, u) \in \mathbb{R}^{k+1}$ are now denoted by (y, z), where $y = (x_1, \dots, x_{k+1})$ and $z = (x_{k+2}, \dots, x_n, u)$. In particular $(y^*, z^*) = (x^*, 0)$, so that (y^*, z^*) belongs to the set of solutions to

$$F(y, z) = \begin{pmatrix} f(x^*) & c_1 & \dots & c_k \end{pmatrix}^T. \tag{19}$$

This determines a C^1-surface locally near (y^*, z^*), since $\frac{\partial F}{\partial y}(y^*, z^*)$ is invertible.

Consequently there exists a C^1-function ψ defined on an open neighbourhood of z^* and (small) closed balls $B_1 = \bar{B}(y^*, \alpha)$, $B_2 = \bar{B}(z^*, \beta)$ such that $\psi\colon B_2 \to B_1$ and that (19) in $B_1 \times B_2$ is solved precisely by the points on the graph of ψ.

In particular (19) is solved by $(\psi(z), z)$ for $z = (x^*_{k+2}, \dots, x^*_n, u)$ with $u \in [-\beta, \beta]$. From the first entry in F it is seen that for *every* $u \in [-\beta, \beta]$ it holds that

$$f(\psi(x^*_{k+2}, \dots, x^*_n, u), x^*_{k+2}, \dots, x^*_n) + u = f(x^*).$$

Since $u \mapsto \psi(x^*_{k+2}, \dots, x^*_n, u)$ is continuous, the pre-image of every ball $B(x^*, r)$ contains an interval $]-\gamma, \gamma[$. On this the function $u \mapsto f(x^*) - u$ is monotone decreasing, so from the equation above it is seen that $B(x^*, r)$ contains points x', x'' for which

$$f(x') > f(x^*), \qquad f(x'') < f(x^*). \tag{20}$$

Because $(\psi(z), x^*_{k+2}, \dots, x^*_n) \in \mathcal{F}$, the number $f(x^*)$ is not an extreme value on $\mathcal{F}$. This proves the theorem.

Remark. In the special case in which the first row of M is a linear combination of the others, it is obvious that there exists certain scalars $\lambda_1,\dots,\lambda_k$ such that $\nabla f(x^*) + \lambda_1 \nabla g_1(x^*) + \dots + \lambda_k \nabla g_k(x^*) = 0$. This implies

$$\nabla(f + \lambda_1 g_1 + \dots + \lambda_k g_k)(x^*) = 0 \tag{21}$$

so the function $f + \lambda_1 g_1 + \dots + \lambda_k g_k$ has a critical point at x^*.

Therefore, if it is known that $g_1,\dots,g_k$ have linearly independent gradients, the extremum points of f under the contraints can according to the above be determined among the critical points of

$$f + \lambda_1 g_1 + \dots + \lambda_k g_k. \tag{22}$$

This clearly gives $n + k$ unknowns and the same number of equations, namely $x_1,\dots,x_n$ and $\lambda_1,\dots,\lambda_k$ that appear the n equations (21) as well in the k constraints.

This method of *constrained optimisation* is often called the method of *Lagrangian multipliers* (these are the numbers $\lambda_1,\dots,\lambda_k$). In practice, however, the theorem is more convenient because linear independence of the gradients need not be verified first. Furthermore, it is often technically easier just to examine the points at which M does not have maximal rank.

Med venlig hilsen
Jon Johnsen

Overview no. 3

2nd gathering, Wednesday 8 September. We meet at 8.15 in G5-109.
In the exercises we shall look at

Local maximum: Determine whether $f(x, y) = x^3 + x^2 + 3x^2y + 3xy^2 + y^3$ has a local extremum at the origin.

Prove that $f(x, y) = y^4 - x^2$ does not have a local maximum at $(0, 0)$.

Explain why the necessary condition provided by Theorem 1, which states that every eigenvalue $\lambda \leq 0$, is *not* sufficient for a maximum.

The integral remainder term: Carry out the induction argument for Taylor's formula in line (1) above.

o**-functions:** Show that if two functions $f(t)$ and $g(t)$ both are $o(t^n)$ for $t \to 0$, then so is $f(t) + g(t)$.

Would it be justified to write: $o(t^n) + o(t^n) = o(t^n)$?

Taylor's limit formula: Use (2) to show that the Taylor polynomial of degree 2 in formula (4) is correct.

Show next that the remainder term in (4) *is* a term with the $o(\|x - x^*\|^2)$ property, as claimed. [*Hint:* One can use (1)].

Constraints: Do exercise 13.5.51 in the calculus book of Edwards and Penney. That is, consider the production of a buoy from steel plates; it should have circular cross section of radius r and height H, while its ends should be conical og height h (much like a pecil sharpened at both ends !). *Minimise* the surface area (and hence the production costs) under the *constraint* that the volume is a given constant V.

In the lectures we first complete the notes above, beginning with a reminder on the implicit function theorem — please review this at home before the lecture!
We also proceed from page 13 in [SS].

Med venlig hilsen
Jon Johnsen

Overview no. 4

Today we covered the rest of the notes above on optimisation of real functions with constraints. These notes have also been expanded accordingly with additional material today.

3rd gathering, Wednesday 15 September. Here we meet in **G5-110** at 8.15. In the lectures we now commence with the calculus of variations, pp. 13-31 in [SS]. Exercises:

Local extrema: Show that $f(x, y) = 4x^2e^y - 2x^4 - e^{4y}$ has 2 critical points; find them. Show that they are *both* local maxima ("Two mountains without a valley"). Surprised? — could this happen in dimension 1 ?

Constraints: Find the points on the surface $xy - z^2 = 1$ that has the least distance to the origin.

Find the dimensions of the largest rectangular box that can be inscribed in the ellipsoid given by

$$(\tfrac{x}{a})^2 + (\tfrac{y}{b})^2 + (\tfrac{z}{c})^2 = 1.$$

Consider a current I that branches into three, so that $I = I_1 + I_2 + I_3$, where each I_j runs through a resistor R_j, where the power dissipation is $R_jI_j^2$. Supposing that the actual sizes of I_1, I_2, I_3 will minimise the total loss of energy, find the ratios I_1/I_2, I_2/I_3 and I_1/I_3.

Old exercises in the remaining time.

Med venlig hilsen
Jon Johnsen

Overview no. 5

Today we covered all essential parts of p. 13–31. The special cases in Section 1.4 is left for you to read.

In particular we covered the deduction of the Euler equation as a necessary condition. However, the precise argument for differentiation under the integral sign (that is, formula (8) in [SS]) is postponed to the next lecture.

4th gathering, Wednesday 22 September. As exercises you can do 1.1.1 at home along with your reading (recommended!). Then continue with 1.2.1+3+4+5+6 from [SS] in the groups.

In the lecture we shall complete the proof for the Euler equation (cf the above) and continue with pp. 35-42 in [SS]. Here we shall meet other, less strict terminal conditions, as well as the *transversality* conditions they give rise to.

5th gathering, Friday 1 October. We continue in the lectures with Section 1.6–7 and some material on convex/concave functions; for the latter we can at least refer to Appendix B in [SS].

In the exercises we look at 1.3.1 and 1.4.1–4 (easy if you have read Section 1.4!). Also 1.5.1 on transversality conditions.

Med venlig hilsen
Jon Johnsen

Overview no. 6

Last time we covered Chapter 1.6 and 1.7, leaving the extension to vector functions in Chapter 1.8 for you to read.

6th gathering, Thursday 7 October. We will finish Chapter 1 in [SS] with some remarks and examples.

In the exercises we will do 1.6.1–2, 1.6.4 and 1.7.4.

7th gathering: First set of mandatory assignments. This consists in both a practical and a theoretical exercise, namely

- Find the dimensions of the largest rectangular box that can be inscribed in the ellipsoid given by

  $$(\tfrac{x}{a})^2 + (\tfrac{y}{b})^2 + (\tfrac{z}{c})^2 = 1.$$

  Determine also the volume of this box. (*Note:* You may assume, without proof, that this box has its sides parallel to the coordinate axes.)

- Solve Exercise 1.8.4 in [SS].

As for the principles of this evaluation: Both exercises should be solved by each of you, individually. You should also sign every sheet of the solution and present it to me no later than Monday 25 October 2010.

Med venlig hilsen
Jon Johnsen

Overview no. 7

8th gathering, Thursday 14 October.

In the exercises we will do 1.7.1–2 and 1.8.1–2.

The lectures will give more examples from and remarks on Section 1.8–9.

Med venlig hilsen
Jon Johnsen