Lecture notes TI-course Math II autumn 2014

Plan course

1. Smooth static optimization: Weierstrass, Fermat, Lagrange; second order conditions, envelope theorem, shadow price.

2. Convex and mixed static optimization: convex Fermat, Karush-Kuhn-Tucker, Fritz-John; convexity, separation, subgradient, shadow price, duality.

3. Dynamic optimization I: the Calculus of Variations (Euler, transversality, Euler-Lagrange), Optimal Control (Pontryagin's maximum principle), Hamiltonian, shadow price, bang-bang, saturation,

4. Dynamic optimization II: Dynamic Programming (Bellman); policies in open and closed loop form; introduction to infinite horizon stationary problems.

5. Dynamic Optimization III: infinite horizon stationary problems.

6. Wrap up of the course.

Novelties

Some of the contents of these notes are novel compared to what you find in textbooks.

1. Lecture on smooth static optimization

   (a) No use is made of second order conditions in the analysis of concrete problems; instead the theorem of Weierstrass is used.

   (b) Emphasis on the clear structure of the theory: there are only four issues—existence, necessary conditions, sufficient conditions, numerical solution—each of which is dominated by its own principle

   (c) No distinction is made in the proof of the theorem of Weierstrass between whether the objective function is bounded or unbounded.

   (d) Short self-contained proof of the Lagrange multiplier rule, using only the theorem of Weierstrass, and not the deep implicit function theorem.

(e) For each optimization method in this course a convincing illustration is given to show that this technique is indispensable. This is not a usual practice. For example, most problems that are given in textbooks to illustrate the multiplier method, can just as well be solved by eliminating constraints and then finding stationary points.

2. Lecture on convex and mixed static optimization

   (a) Efficient solution of the facility location problem of Fermat-Weber by means of the convex Fermat theorem. The usual textbook solution by Fermat's theorem does not cover all cases.

   (b) Formulation of the Karush Kuhn Tucker conditions in terms of subdifferentials. This recent formulation by Vladimir Protassov is simpler to memorize and to use; some previously 'unsolvable' problems can be solved, using this formulation.

   (c) Formulation of the Fritz John conditions in terms of convex hulls of gradients. This result appears for the first time in these notes. It runs parallel to the result of Vladimir Protassov for KKT.

   (d) A proof of the theorem of Minkowski on polyhedra by means of the Fritz John conditions. This result is recent work by Vladimir Protassov. The proof of this theorem that is usually given in textbooks is not correct; moreover, this proof provides a convincing illustation of the advantage of FJ over Lagrange. Such examples are rare; I know of only one other such example: the analysis of the shortest distance problem of Apollonius for a point and a conc section (circle, ellipse, parabola, hyperbola).

3. Lecture on dynamic optimization I

   (a) The proof that the Euler equation is a necessary condition that is given does not require the technical lemma of du Bois Reymond. This proof is due to Vladimir Tikhomirov.

4. Lecture on dynamic optimization II

   (a) A simple proof that the Bellman equation gives a criterion for optimality provided that the value function is C^1; it is a combination of a reduction to the case of an objective function that depends only on the endstate and a simple observation 'about a boat journey'.

   (b) A powerful method that combines the advantages of Euler/Pontryagin with these of Bellman, and leads to a complete analysis.

   (c) A simplification of the proof that the Bellman equation gives a criterion for optimality for infinite horizon stationary problems in discrete time.

5. Lecture on optimization III

(a) Simplified self-contained account of the theory of infinite horizon stationary problems in discrete time.

1 Lecture on smooth static optimization: Weierstrass, Fermat and Lagrange

1.1 Introduction

Four step method. All extremal problems that can be solved analytically at all, can be solved in a systematic way, in four steps: 1) establish existence (using the theorem of Weierstrass), 2) write the first order necessary conditions (using the theorems of Fermat, Lagrange, ...), 3) analyze the conditions, 4) write the conclusion. Thanks to this systematic method, you 'do not have to be a Newton' if you want to solve extremal problems; everybody can learn the craft to bulldozer over all extremal problems by one method only. Of course, individual problems can sometimes be solved slightly quicker: by using all sorts of special tricks, but it is an advantage that this is not necessary. **Recommendation: always use the four step method and never use any special tricks.**

Intuition four step method. The allegory of a murder investigation: step 1: we have to make sure that there exist someone who has committed a murder; step 2: we have to produce the means to find the suspects; step 3: we have to make a complete list of all suspects, and then eliminate suspects until only one is left; step 4: present the conclusion of the investigation.

Clear structure theory. It is good to keep in mind that the theory of optimization has a clear structure (however, the details of this clear theory fall outside the scope of this course): it consists of four parts, each governed by its own principle: (1) existence: the principle of compactness, (2) necessary conditions for regular smooth-convex problems: the principle of Fermat-Lagrange, (3) sufficient conditions for regular smooth-convex problems: the principle of embedding in a family of problems, (3) existence: the principle of compactness, (4) numerical solution: the bisection method and its generalization to many variables—the center of gravity method.

1.2 Weierstrass

Value of a problem. For each subset $S \subseteq \mathbb{R}$, one defines its infimum to be the largest lower bound of S—$\inf S = \max\{r \in [-\infty, \infty] | r \le s \, \forall s \in S\}$. S has a minimum iff $\inf S \in S$; then $\min S = \inf S$. The value of an extremal problem $f(x) \to \min, x \in A$, where A is a set and $f : A \to \mathbb{R}$ is a function, is defined to be $v(P) = \inf\{f(a)|a \in A\}$. An admissible point x of the problem, that is, $x \in A$, is a solution of the problem iff $f(x) = v(P)$. There are three different reasons why a problem can have no solution: (1) $v(P) = +\infty$, that is, the problem has no admissible points, (2) $v(P) = -\infty$, that is, the objective function is unbounded below, (3) $v(P) \in \mathbb{R}$, that is, the value can be approximated arbitrarily closely, but it is not assumed by the objective function.

Numerical examples. (1) $x^2 + y^2 \to \min, x, y \in \mathbb{R}, 2x + y^2 + 1 < 0, x^2 + 2y + 1 < 0$, (2) $x^2 - 4xy +$

$3y^2 \to \min, x, y \in \mathbb{R}$, (3) $x^2 + (xy - 1)^2 \to \min, x, y \in \mathbb{R}$.

Intuition value problem. Each problem to find the smallest number of a given finite nonempty set of values has a solution. For infinite sets, the best you can do is define a sort of 'surrogate solution', called the infimum of the set. The great advantage of this concept is that it is always defined. It allows you two break the tough task of finding a smallest number of a collection of values into to more manageable tasks: first calculate the infimum, then check whether the infimum belongs to the collection of numbers: if so, then it is the minimum, if not, then there exists no minimum.

Theorem of Weierstrass. *Ingredients: $f : C \to \mathbb{R}$, $C \subset \mathbb{R}^n$. Assumptions: f continuous; $C \neq \emptyset$; C is closed (for each convergent sequence $\{c_i\}$ in C, its limit $\lim_{i\to\infty} c_i$ belongs to C as well), C is bounded (there exists a number $R > 0$ such that $|x_i| \leq R$ for all $x \in C$ and all $i \in \{1, \ldots, n\}$). Conclusion: existence of a global minimum $\widehat{x}$ of the problem $f(x) \to \min, x \in C$.*

Completeness. We need the fundamental property 'completeness' of the real numbers. One of the many equivalent ways of expressing this is: for each nested sequence of closed intervals $[a_1, b_1] \supseteq [a_2, b_2] \supseteq \cdots$ for which the lengths of the intervals tend to zero, $\lim_{k\to\infty}(b_k - a_k) = 0$, these intervals have a unique real number in common.

Proposition. *Each bounded sequence of real numbers $\{c_k\}$ has a convergent subsequence.*

Proof proposition. Take a closed interval $[A, B]$ that contains the sequence $\{c_k\}$ and let C be the midpoint of the interval. Then at least one of the two subintervals $[A, C]$ and $[C, B]$ contains infinitely many points of the sequence $\{c_k\}$, say $[A, C]$. Then $[A, B] \supseteq [A, C]$, both contain infinitely many points of the sequence $\{c_k\}$ and the length of $[A, C]$ is half the length of $[A, B]$. Continuing in this way one gets a nested sequence of closed intervals with length tending to 0 for which each one of these intervals contain infinitely many points of the sequence $\{c_k\}$. Then it follows readily that the unique common point of this nested sequence of intervals is the limit of a suitable subsequence of $\{c_k\}$.

Corollary. *Each bounded sequence $\{c_k\}$ in $\mathbb{R}^n$ has a convergent subsequence.*

Proof corollary. Choose a subsequence for which the first coordinates converge; this is possible by the proposition. Then take a subsequence of this subsequence for which the second coordinates converge, using the proposition again; then the first coordinates still converge. Do this n times and then you have a convergent subsequence.

Proof theorem of Weierstrass. Choose a sequence $\{c_i\}$ in C with $\lim_{i\to\infty} f(c_i) = \inf f$, the infimum of f (this choice can be made: indeed, if $\inf f \in \mathbb{R}$, then we can choose c_i such that $f(c_i) < \inf f + \frac{1}{i}$ for all i; otherwise we can choose c_i such that $f(c_i) < -i$ for all i). Then choose a convergent subsequence $\{c_{i_k}\}$. Write $\bar{c}$ for the limit. For this subsequence, the limit of the f-values, $\lim_{k\to\infty} f(c_{i_k})$, is still $\inf f$. Then $f(\bar{c}) = \lim_{k\to\infty} f(c_{i_k}) = \inf f$ (the first equality holds as f is continuous). This means that $\bar{c}$ is a solution of the given problem.

For each result of optimization theory we try to give convincing evidence that it is indispensable.

Convincing illustration. The proof of the fundamental theorem of algebra by means of the theorem of Weierstrass (see the proof of theorem 2.8).

Handy consequences of Weierstrass. Same conclusion in the following cases where closedness/boundedness are not satisfied: 1) $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) \to \infty$ if $|x| \to \infty$—we will always use the notation $|x| = \sqrt{\sum_{k=1}^n x_k^2}$— (this is the case if for example $f(x) = x^T Ax + ax + \alpha$ with A a symmetric positive definite $n \times n$-matrix, a an n-row vector and α a number), 2) $f : (a, b) \to \mathbb{R}$ with $f(x) \to \infty$ if $x \downarrow a$ and $x \uparrow b$.

Intuition Weierstrass. Most—maybe all—major results from static optimization (Lagrange multiplier rule, Karush Kuhn Tucker), calculus (implicit function theorem, separation theorem of convex sets), matrix algebra (fundamental theorem of algebra, orthogonal diagonalization of symmetric matrices) can be derived from the theorem of Weierstrass. It is an existence theorem: it establishes that something exists without giving information where this something can be found. This sounds somewhat vague, but in reality it is an invaluable result: it is like a golden key that opens many doors!

1.3 Fermat

Approximation definition derivative. Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $\widehat{x} \in \mathbb{R}^n$. If one can split up the function $\mathbb{R}^n \to \mathbb{R}$ given by $h \mapsto f(\widehat{x} + h) - f(\widehat{x})$ as the sum of a linear function $h \mapsto ah$ for some $a \in (\mathbb{R}^n)^T$ (here ah is the matrix product of the row a and the column h) and some function $r(h)$ that is negligible in the sense that even if we divide it by $|h|$, it still tends to zero for $|h| \to 0$, then we call f differentiable in $\widehat{x}$, we call a the derivative of f in $\widehat{x}$ and we write $f'(\widehat{x}) = a$. We indicate that $r(h)$ is negligible by means of the small Landau o notation: $r(h) = o(|h|)$; so this means that $r(h)/|h|$ tends to zero if $|h|$ tends to zero. One can prove that $f'(\widehat{x})$ equals the gradient $(\frac{\partial f}{\partial x_j}(\widehat{x})) \in (\mathbb{R}^n)^T$ if all partial derivatives $\frac{\partial f}{\partial x_j}$, $1 \leq j \leq n$ of f exist and are continuous.

Intuition approximation definition derivative. This definition of the derivative makes good use of vector concepts and is therefore essentially the same for functions $f(x_1, \ldots, x_n)$ as for functions of one variable; it can also be used in dynamic optimization, where we want to define the derivative of a function $J(x(\cdot))$ where $x(\cdot)$ runs over a space of functions. To calculate derivatives, one should never use the definition, but one should use calculus rules. To begin with one should then use the expression of the derivative in terms of partial derivatives, which one can easily calculate of course. The reason for being interested in derivatives is linearization: we want to reduce nonlinear (and so hard) problems to linear (and so easy) problems. If one is interested in the marginal behavior of a function f near a point $\widehat{x}$, one can linearize, that is, we can approximate the function $x \mapsto f(x)$ by its linearization $\widehat{x} + h \mapsto \widehat{x} + f'(\widehat{x})h$.

Illustrations. (1) $(\widehat{x} + h)^2 - \widehat{x}^2 = 2\widehat{x}h + h^2$; $h \to 2\widehat{x}h$ is linear and $h \to h^2$ is negligible, so the derivative of $x \to x^2$ in $\widehat{x}$ is $2\widehat{x}$, as expected, (2) Problem 2.2.1.

Theorem (Fermat). *Ingredients: $\varepsilon > 0$, $\widehat{x} \in \mathbb{R}^n$, $f : \{x \in \mathbb{R}^n \| x - \widehat{x}| < \varepsilon\} \to \mathbb{R}$. Assumptions: f differentiable in $\widehat{x}$, and $\widehat{x}$ is a local minimum for the problem $f(x) \to \min, x \in \mathbb{R}^n, |x - \widehat{x}| < \varepsilon$ (that is, it is a global minimum in $\widehat{x}$ on $\{x \in \mathbb{R}^n \| x - \widehat{x}| < \bar{\varepsilon}\}$ for $\bar{\varepsilon} > 0$ sufficiently small). Conclusion: $f'(\widehat{x}) = 0$.*

Proof. We argue by contradiction. Assume $f'(\widehat{x}) \neq 0$. Consider the vector $v = \frac{f'(\widehat{x})^T}{|f'(\widehat{x})|^2}$. We have, by the definition of the derivative, $f(\widehat{x} - \alpha v) - f(\widehat{x}) = f'(\widehat{x})(-\alpha v) + r(-\alpha v)$. Then we apply the rule $w^T w = |w|^2$ to $w = f'(\widehat{x})$ and this gives the outcome $-\alpha + r(-\alpha v) = \alpha(-1 + \frac{r(-\alpha v)}{\alpha})$. This shows, as $\lim_{\alpha \downarrow 0} \frac{r(-\alpha v)}{\alpha} = 0$, that for $\alpha > 0$ sufficiently small we get $f(\widehat{x} - \alpha v) - f(\widehat{x}) < 0$. This contradicts the local minimality of $\widehat{x}$, as required.

Numerical illustration

1. Modeling and existence. $f(x) = x^2 - 4|x| + 5 \to \min, x \in [-1, 10]$; A solution $\widehat{x}$ exists by Weierstrass.

2. Necessary condition. Fermat: $0 = f'(x) = 2x - 4x/|x|$.

3. Analysis. Candidate solutions: (1) stationary points: $x > 0$ gives $2x - 4 = 0$ so $x = 2$, $x < 0$ gives $2x + 4 = 0$ so $x = -2$, which is not in $[-1, 10]$ (2) points of non-differentiability: $x = 0$, (3) boundary points $-1, 10$. Comparison candidate solutions: $f(2) = 1, f(0) = 5, f(-1) = 2, f(10) = 65$: f has smallest value in 2.

4. Conclusion. The problem has a unique solution: $\widehat{x} = 2$.

Convincing illustration. The optimal pipeline from source to coast for transport oil to island can be determined by means of the theorem of Fermat (section 1.4.6).

Intuition Fermat. In an optimum, a marginal change has a negligible effect. A rational agent acts in the margin and he will only stop when marginal changes have a negligible effect. The proof of Fermat can be viewed in this way. If $f'(\widehat{x}) \neq 0$ for some point $\widehat{x}$, then a small change in the opposite of the direction of the gradient, $-\nabla f(\widehat{x}) = -f'(\widehat{x})^T$, gives a linear decrease of the f-value, so such a point $\widehat{x}$ cannot be a minimum. In other words, in a minimum $\widehat{x}$ one has $f'(\widehat{x}) = 0$.

1.4 Lagrange

Theorem (Lagrange multiplier rule). *Ingredients:* $F = (f_0, \ldots, f_m)^T : \mathbb{R}^n \to \mathbb{R}^{m+1}$, $\widehat{x} \in \mathbb{R}^n$. *Assumptions:* F *continuously differentiable in* $\widehat{x}$, *and* $\widehat{x}$ *is a local minimum of* $f_0(x) \to \min, f_j(x) = 0, 1 \le j \le m$. *Conclusion: there exist numbers* $\lambda_i, 0 \le i \le m$, *not all zero, such that* $\sum_{i=0}^m \lambda_i f_i'(\widehat{x}) = 0$.

Additional terminology and notation. If the vectors $f_i'(\widehat{x}), 1 \le i \le m$, are linearly independent, then $\lambda_0 \ne 0$, and so we may take $\lambda_0 = 1$, without loss of generality. Then λ_i is called the Lagrange multiplier of the constraint $f_i(x) = 0$ for all $i \in \{1, \ldots, m\}$. This can also be written as: $\widehat{x}$ is a stationary point for the Lagrange function $\mathcal{L}(x) = \sum_{i=0}^m \lambda_i f_i(x)$ for $\lambda_0 = 1$ and a suitable choice of multipliers $\lambda_i, 1 \le i \le m$. Warning: sometimes the Lagrange function for $\lambda_0 = 1$ is defined to be $f_0(x) - \lambda_1 f_1(x) - \cdots - \lambda_m f_m(x)$; then one gets the opposite numbers for the multipliers.

Definition derivative vector function. Given a vector function $f : \mathbb{R}^n \to \mathbb{R}^m$ and a point $\widehat{x} \in \mathbb{R}^n$. If one can split up the function $\mathbb{R}^n \to \mathbb{R}^m$ given by $h \mapsto f(\widehat{x}+h) - f(\widehat{x})$ as the sum of a linear function $h \mapsto Ah$ for some $m \times n$-matrix A and some function $h \mapsto r(h)$ that is negligible in the sense that even if we divide it by $|h|$, it still tends to zero for $|h| \to 0$, then we call f differentiable in $\widehat{x}$, we call A the derivative of f in $\widehat{x}$ and we write $f'(\widehat{x}) = A$. We indicate that $r(h)$ is negligible by means of the small Landau o notation: $r(h) = o(|h|)$; so this means that $r(h)/|h|$ tends to zero if $|h|$ tends to zero. One can prove that $f'(\widehat{x})$ equals the $m \times n$-Jacobi matrix $(\frac{\partial f_i}{\partial x_j}(\widehat{x}))_{ij}$ if all partial derivatives of f exist and are continuous.

***Proof Lagrange multiplier rule.** For simplicity we give the proof in a characteristic special case: $m = 1, n = 2$, so we consider the problem $f_0(x_1, x_2) \to \min, f_1(x_1, x_2) = 0$. We proceed by contradiction. Assume that the two vectors $f_i'(\widehat{x}) = (\frac{\partial f_i}{\partial x_1}(\widehat{x}), \frac{\partial f_i}{\partial x_2}(\widehat{x})), i = 1, 2$ are linearly independent in $(\mathbb{R}^2)^T$. Then the 2×2-matrix $F'(\widehat{x}) = (\frac{\partial f_i}{\partial x_j}(\widehat{x}))_{ij}$ is invertible. We assume that $\widehat{x} = 0$ and $f_0(\widehat{x}) = 0$ (and so that $F(\widehat{x}) = 0$), as we may without loss of generality: this can be achieved by replacing x by $x - \widehat{x}$, and $f_0(x)$ by $f_0(x) - f_0(\widehat{x})$.

Now we make some choices. We choose $\delta > 0$ arbitrary but so small that : (1) $|x| \le \delta \Rightarrow F'(x)$ is invertible; this is possible as $F'(0)$ is invertible and by changing an invertible matrix slightly it remains invertible; (2) $|x| = \delta \Rightarrow |F(x) - F'(0)x| < \frac{1}{2}a|x|$ with $a = \min_{|x|=1} |F'(0)x| > 0$ (this minimum is assumed by Weierstrass and it is > 0 as $F'(0)$ is invertible); this is possible as $F(x) = F'(0)x + o(|x|)$ ('F differentiable in 0'). Then we choose a positive number $b < b(\delta) = \frac{1}{2}\min_{|x|=\delta} |F(x)|$ (this minimum is assumed by Weierstrass, and it is > 0 as $|x| = \delta \Rightarrow |F(x)| \ge |F'(0)x| - |F(x) - F'(0)x| \ge a|x| - \frac{1}{2}a|x| > 0$; $\widehat{x} = 0$ is a global minimum of $f_0(x) \to \min, f_1(x) = 0, |x| < \delta$.

After these preliminaries, we can give the proof. We consider the auxiliary problem $|F(x) + (b, 0)^T| \to \min, |x| \le \delta$. This problem has a solution $\bar{x}$ by Weierstrass. It is impossible that $|\bar{x}| = \delta$, as this gives $|F(\bar{x}) + (b, 0)^T| \ge |F(\bar{x})| - b > b = |F(0) + (b, 0)^T|$, in contradiction with the minimality of $\bar{x}$ for the auxiliary problem. Therefore, $|\bar{x}| < \delta$. It is also impossible that $F(\bar{x}) + (b, 0)^T \ne (0, 0)$, as this gives, with the vector h for which $F'(\bar{x})h = -(F(\bar{x}) + (b, 0)^T)$ (here we use (1)) that the vector $F(\bar{x} + th) + (b, 0)^T$, with $t > 0$ is equal to $F(\bar{x}) + F'(\bar{x})(th) + r((th) + (b, 0)^T$ with $r(th) = o(t)$—by the differentiablity of F in $\bar{x}$—and this is, by the choice of h equal to $(1-t)(F(\bar{x})+(b,0)^T)+o(|t|)$, and so this vector is readily seen to be shorter than $F(\bar{x})+(b,0)^T$ for $t>0$ sufficiently small. This would contradict the minimality property of $\bar{x}$. Therefore, $F(\bar{x})+(b,0)^T=(0,0)$, that is, $f_0(\bar{x})=-b, f_1(\bar{x})=0$. Moreover, we recall that $|\bar{x}|<\delta$. This contradicts (3).

Hint for use. In each application, you should begin by excluding the bad case $\lambda_0=0$.

Numerical illustration. How to maximize your utility $3\ln x+2\ln y$ of spending 5 euros on playing games in a game hall (50 euro cents for one game) and eating icecreams (which cost 1 euro each).

1. Modeling and existence. $f_0(x)=-3\ln x_1-2\ln x_2\to\min, f_1(x)=\frac{1}{2}x_1+x_2-5=0, x_1,x_2>0$; existence of a solution $\widehat{x}$ follows from Weierstrass: elimination of x_2 gives $-3\ln x_1-2\ln(5-\frac{1}{2}x_1)$ and this tends to ∞ for $x\downarrow 0$ and $x\uparrow 10$.
2. Necessary conditions. Lagrange function: $\mathcal{L}=\lambda_0(-3\ln x_1-2\ln x_2)+\lambda_1(\frac{1}{2}x_1+x_2-5)$.

   Lagrange: $0=\mathcal{L}_x$, that is, $\mathcal{L}_{x_1}=\frac{3\lambda_0}{x_1}+\frac{\lambda_1}{2}=0, \mathcal{L}_{x_2}=\frac{2\lambda_0}{x_2}+\lambda_1=0$.

   Exclusion bad case: $\lambda_0=0$ implies $\lambda_1=0$, which contradicts the requirement that at least one of the multipliers is nonzero.

   Put $\lambda_0=1$.
3. Analysis. Eliminate λ_1: $x_1=3x_2$. Use the constraint and then you get $x=(6,2)^T$
4. Conclusion. There is a unique solution: it is optimal to play six games and to eat two ice creams.

Numerical illustration.

1. Modeling and existence. $f(x,y)=5x^2+2xy+3y^2\to\min, g(x,y)=7x^2+2xy+4y^2-3=0$; existence of a solution $\widehat{x}$ follows from Weierstrassas as we can write $g(x,y)=\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}7 & 1\\ 1 & 4\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$ and the matrix $\begin{pmatrix}7 & 1\\ 1 & 4\end{pmatrix}$ turns out to be positive definite.
2. Necessary conditions. Lagrange function $\mathcal{L}=\lambda_0(5x^2+2xy+3y^2)+\lambda_1(7x^2+2xy+4y^2-3)$.

   Lagrange: $0=\mathcal{L}'(x,y)$, that is, $\mathcal{L}_x=\lambda_0(10x+2y)+\lambda_1(14x+2y)=0, \mathcal{L}_y=\lambda_0(6y+2x)+\lambda_1(8y+2x)=0$. Exclusion bad case: $\lambda_0=0\Rightarrow\lambda_1\neq 0, 14x+2y=0, 8y+2x=0\Rightarrow x=y=0$; this contradicts the constraint.
3. Analysis. Eliminate λ: $\frac{10x+2y}{14x+2y}=\frac{6y+2x}{8y+2x}\Rightarrow\frac{y}{x}=-1$ or 2. Substitute this in $g(x,y)=0$. Then we get some suspects, a comparison of f-values shows that $(-\frac{1}{3},\frac{1}{3})$ and $(\frac{1}{3},-\frac{1}{3})$ are the minima.
4. Conclusion. There are two minima, $(-\frac{1}{3},\frac{1}{3})$ and $(\frac{1}{3},-\frac{1}{3})$

Convincing illustrations. (1) The hinging quadrangle (section 3.3.6); (2) Diagonalizability of a symmetric matrix (Problem 3.3.7).

Intuition Lagrange. In an optimum an *admissible* change dx has a negligible effect dy on the f-value: that is, $f_i'(x)dx = 0, i = 1, \ldots, m \Rightarrow f_0'(x)dx = 0$. By the theory of linear equations, this is equivalent to: $f_0'(x)$ is a linear combination of the $f_i'(x), i = 1, \ldots, m$, that is, to the Lagrange equations. The secret of the power of the Lagrange multiplier rule is the reversal of the order of the two natural tasks that have to be performed: elimination and differentiation. By this reversal, the hardest task, elimination, is turned from a nonlinear problem into a linear problem (see section (3.5.1) in [B-T] for details).

1.5 Second order conditions

Definiteness for symmetric matrices. We recall some definitions from linear algebra that are needed to formulate the first and second order conditions for local minimality.. An $n \times n$-matrix A is called symmetric if $A^T = A$. A symmetric matrix A is called positive semi-definite if $x^T Ax \geq 0$ for all (nonzero) $x \in \mathbb{R}^n$. A symmetric matrix is called positive definite if $x^T Ax > 0$ for all nonzero $x \in \mathbb{R}^n$. The nullspace of a matrix B is the solution space of $Bx = 0$. A symmetric $n \times n$-matrix A is positive semi-definite on the nullspace of an $m \times n$-matrix B if the following implication holds: $Bx = 0 \Rightarrow x^T Ax \geq 0$. A symmetric $n \times n$-matrix A is positive definite on the nullspace of an $m \times n$-matrix B if the following implication holds: $Bx = 0, x \neq 0 \Rightarrow x^T Ax > 0$.

Illustration. There is an efficient method to check whether a given symmetric matrix is positive definite or positive semi-definite (see section (5.2.2) in [B-T]).

Definition second derivative function of several variables. The most natural definition of the second derivative is as the derivative of the derivative. For functions $f : \mathbb{R}^n \to \mathbb{R}$, the derivative $f'(x)$ is an n-dimensional row vector. If f is C^1, this is the row of all partial derivatives $\frac{\partial f}{\partial x_j}$. Taking again the derivative, one gets the second derivative $f''(x)$, the $n \times n$-matrix for which the i-th row is the derivative of $f'(x)$ with respect to x_i. If f is C^2—that is, all partial derivatives of order two of f exist and are continuous—this is the row of second order partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}, i = 1, \ldots, n$. That is, $f''(x)$ is the hessian $(\frac{\partial^2 f}{\partial x_i \partial x_j})_{ij}$, which is a symmetric matrix as the order in which partial derivatives are taken does not matter for C^2-functions. One can show that $f(\widehat{x} + h) - f(\widehat{x}) = f'(\widehat{x})h + \frac{1}{2}h^T f''(\widehat{x})h + o(|h|^2)$, where $o(|h|^2)$ denotes a function of h that after division by $|h|^2$ still tends to zero if $|h| \to 0$.

Theorem (first and second order conditions for local minimality)

1. *Unconstrained problems: consider the point $\widehat{x}$ and the problem $f(x) \to \min, x \in \mathbb{R}^n$ with $f : \mathbb{R}^n \to \mathbb{R}$ a C^2-function (all second order partial derivatives exist and are continuous): (1)*

necessary conditions for local minimality of $\widehat{x}$: $f'(\widehat{x}) = 0$ and $f''(\widehat{x})$ is positive semi-definite; (2) sufficient conditions for local minimality of $\widehat{x}$: $f'(\widehat{x}) = 0$ and $f''(\widehat{x})$ is positive definite.

2. *Equality constrained problems: consider problem $f_0(x) \to \min, G(x) = 0$, with $G : \mathbb{R}^n \to \mathbb{R}^m$ a C^2 function and an admissible point $\widehat{x} \in \mathbb{R}^n$ such that $G'(\widehat{x})$ has rank m: (1) necessary conditions for local minimality of $\widehat{x}$: $\mathcal{L}_x(\widehat{x}, \lambda) = 0, \mathcal{L}_{xx}(\widehat{x}, \lambda)$ is positive semidefinite on the nullspace of $G'(\widehat{x})$ for a suitable choice of $\lambda = (\lambda_0, \bar{\lambda})$ with $\lambda_0 = 1$; (2) sufficient conditions for local minimality of $\widehat{x}$: $\mathcal{L}_x(\widehat{x}, \lambda) = 0, \mathcal{L}_{xx}(\widehat{x}, \lambda)$ is positive definite on the nullspace of $G'(\widehat{x})$ for a suitable choice of $\lambda = (\lambda_0, \bar{\lambda})$ with $\lambda_0 = 1$.*

On the proof. We only give the derivation of the second order necessary conditions for unconstrained problems. One has $f(\widehat{x} + h) - f(\widehat{x}) = f'(\widehat{x})h + \frac{1}{2}h^T f''(\widehat{x})h + o(|h|^2)$ by the definition of the second order derivative. By Fermat one has $f'(\widehat{x}) = 0$. Substituting this, dividing by $|h|^2$ and taking the limit where h runs in a straight line to the origin, $tu, t \downarrow 0$ with u a nonzero vector, we get that $\lim_{t\downarrow 0} \frac{f(\widehat{x}+tu)-f(\widehat{x})}{t^2} = u^T f''(\widehat{x})u$; so by local minimality of $\widehat{x}$, we get $u^T f''(\widehat{x})u \geq 0$; as u is an arbitrary nonzero vector, it follows that the matrix $f''(\widehat{x})$ is positive semidefinite.

What is the point of second order conditions? These conditions give insight but are not to be used in the solution of concrete problems. The insight is that the first and second order necessary conditions are almost sufficient for local optimality. **It is recommended not to use second order conditions in the analysis of concrete problems**, as this requires cumbersome computations of minors of matrices. It is better to use the theorem of Weierstrass to complement the analysis by first order necessary conditions. This is what is done in the four step method.

1.6 Envelope theorem and shadow price interpretation multipliers

Envelope theorem. *Consider the family of problems $(P_y) : f(x, y) \to \min, x \in \mathbb{R}^n$, where the parameter y runs over $\mathbb{R}^m$. Here $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is assumed to be C^2. Assume that (P_y) has a unique global minimum $x(y)$ for all y in a neighborhood of the origin in $\mathbb{R}^m$. Assume that the first and second order sufficient conditions hold in a point $\widehat{x}$ for the problem (P_0). Then $S'(0) = f_y(\widehat{x}, 0)$, where S is the optimal value function of the family of problems (P_y) and $\widehat{x}$ is the solution of (P_0).*

Remark. This result can be read as follows: one can differentiate $S(y) = f(x(y), y)$ in $y = 0$ just as if $x(y)$ would not depend on y.

Sketch proof. We have $S(y) = f(x(y), y)$ for y in a neighborhood of the origin. The chain rule gives: $S'(y) = f_x(x(y), y)x'(y) + f_y(x(y), y)$ (that $x(y)$ is differentiable, follows from the implicit function theorem). Applying Fermat to the minimality property of $x(y)$ gives that $f_x(x(y), y) = 0$. Substituting this and taking $y = 0$ gives the statement of the theorem.

Intuition envelope theorem. At first sight one might expect that in order to determine the sensitivity of the optimal value for small changes in a parameter in the problem, you have to solve the problem after small changes of the parameter. However this is not necessary by virtue of the envelope theorem.

Theorem (Shadow-price interpretation multiplier (=envelope theorem for equality constrained problems)). *Consider the family of problems (P_y) : $f_0(x) \to \min, G(x) = y$ with $f_0 : \mathbb{R}^n \to \mathbb{R}$ and $G : \mathbb{R}^n \to \mathbb{R}^m$ a C^2 function and an admissible point $\widehat{x} \in \mathbb{R}^n$ such that $G'(\widehat{x})$ has rank m, and where the parameter runs over $\mathbb{R}^m$. Assume that (P_y) has a unique global minimum $x(y)$ for all y in a neighborhood of the origin in $\mathbb{R}^m$. Then $S'(0) = -\bar{\lambda}$, where S is the optimal value function of the family of problems (P_y) and $\widehat{x}$ is the solution of (P_0), and where $\lambda = (\lambda_0, \bar{\lambda})$ is the Lagrange multiplier for which $\lambda_0 = 1$.*

Proof. For simplicity we only consider the case $m = 1$. We have $S(y) = f_0(x(y))$ and so, after differentiation, $G'(x(y))x'(y) = 1$. By applying the Lagrange multiplier rule to the problem (P_y) and the point $x(y)$ we get the Lagrange equations $f_o'(x(y)) + \lambda(y)G'(x(y)) = 0$. Therefore,

$$S'(y) = -\lambda(y)G'(x(y))x'(y) == -\lambda(y).$$

In particular $S'(0) = -\lambda(0) = -\bar{\lambda}$ as required.

Illustration. How does the optimal value of the Fermat-Weber facility location problem change if we change the position of one of the three points? Solution. Assume that all three angles of the triangle formed by the three given points are $< 2\pi/3$. Then the solution is the Torricelli point $\widehat{x}$. Then the optimal value of the problem is $S(a_1, a_2, a_3) = \sum_{i=1}^{3} |\widehat{x} - a_i|$; of course $\widehat{x}$ depends on a_1, a_2 and a_3. Therefore, the task to compute the derivative of $S(a_1, a_2, a_3)$ with respect to a_1, might look daunting at first sight. However, we may ignore, by the envelope theorem, the dependence of $\widehat{x}$ on a_1. This makes everything very simple. We get that the gradient of the optimal value function $S(a_1, a_2, a_3)$ with respect to a_1 equals $\frac{a_1 - \widehat{x}}{|a_1 - \widehat{x}|}$, that is, it is the unit vector that has the same direction as the vector with initial point the Torriccelli point and final point a_1.

Intuition envelope theorem/shadow price interpretation multiplier. Again, at first sight one might expect that in order to determine the sensitivity of the optimal value for small changes in a parameter in the problem, you have to solve the problem after small changes of the parameter. However this is not necessary by virtue of this envelope theorem. Moreover, this result shoes that the Lagrange multipliers are more than just auxiliary numbers that are used to solve extremal problems with equality constraints: these numbers have a valuable interpretation: as a measure how sensitive the optimal value of a problem is for changing a constraint $f_i(x) = 0$ into a constraint $f_i(x) = y_i$ for some number y_i close to zero.

1.7 Problems on the lecture notes:

1. Explain why the following simplification of the four step method is wrong: skip the verification of the existence of a global minimum; if the necessary conditions lead to some candidates, and one of these has a lower value of the objective function than the others, then this is guaranteed to be the unique global minimum of the problem.

2. How do you check for concrete optimization problems that the admissible set is closed? [Hint: appeal to a general result on a system of equalities and inequalities.]

3. Give the details of the proof that a bounded sequence of real numbers has a convergent subsequence. Where is this result needed in the proofs of the underlying results of the four step method?

4. Derive from the theorem of Weierstrass the first handy consequence of Weierstrass given above.

5. Derive from the theorem of Weierstrass the second handy consequence of Weierstrass given above.

6. *In the proof of the multiplier rule, supply the details of the proof that for $\delta > 0$ sufficiently small, property (1) holds true.

7. *In the proof of the multiplier rule, supply the details of the proof that $F(\bar{x}) + (b, 0)^T \neq 0$ is impossible.

8. Second order conditions are not used in the four step method. So what is the point of these conditions?

9. Explain the terminology 'shadow price interpretation'.

10. *Give a precise proof of the envelope theorem for $m = n = 1$; in particular, prove that $x(y)$ is differentiable, using the implicit function theorem (look this theorem up in your Analysis textbook for Math I, for example).

1.8 Your tasks between lecture 1 and lecture 2.

You are expected to do as follows.

- In groups of TWO you have to solve the ten problems on the lecture notes given above, as well as the following ten optimization problems by means of the four step method, and to hand in your work on 17-9-2013. Present complete solutions (not just some calculations). The optimization problems are from [B-T] Appendix D:

  D1.4, D1.5,

  D2.1, D2.2, D2.4, D2.5,

  D3.2, D3.3, D3.4, D3.5.

- Study these lecture notes carefully (if something is not crystal clear or left undefined or not explained in this compactly written text, look it up in [B-T]).
- Study carefully in [B-T] section 0.2 Lunch, Dinner and Dessert pp xiv-xvi and all royal roads: as described in p6 (ch1), p87 (ch2), p138 (ch 3), p.262 (ch5).
- Look through the sections in ch1-3 of [B-T] with applications.
- Study the texts about shadow price interpretations of Lagrange multipliers p.185, 186 and p. 244, 245.
- Look up in the Appendices what you need (for example on the theorem of Weierstrass and on matrix algebra).
- Try to look through the chapters 1-3 in [B-T].
- Make a brave effort to come to grips with the advanced texts and problems; these are in scriptsize and indicated by a star *.

2 Lecture on convex static optimization: Karush Kuhn Tucker, Fritz John

2.1 Convex sets and functions

Intuition on the interest of convexity. Linearization is a powerful technique. Therefore, it is of great interest that there is, besides the class of differentiable functions, still another class of functions for which we can realize the linearization idea. This is the class of convex functions. For each convex function f and each point $\widehat{x}$, one can approximate the function $x \mapsto f(x)$ by a function of the form $\widehat{x} + h \mapsto f(\widehat{x}) + ah$ for some rowvector a. This can be viewed as a linearization. We have to pay a price for this achievement: a need not be unique. For example, take the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$ (the absolute value function) and $\widehat{x} = 0$. Then we can take for a each number in the segment $[-1, 1]$. This is a disadvantage, but on the other hand optimization problems that are 'made' with convex functions have great advantages: for example, the first order necessary conditions are necessary *and sufficient*. Thus, one gets a criterion for optimality. Moreover, there is no distinction between local and global minima: each local minimum is automatically a global minimum. Another advantage is that convex problems can be solved numerically in an efficient way, and with a certificate of quality (however numerical solution of optimization problems falls outside the scope of this course).

***On mixed smooth-convex problems.** It is a natural question whether one can give moreover necessary problems for mixed smooth-convex problems, that is, problems for which the ingredients are partly smooth and partly convex. The answer is affirmative. We give the best-known example: we consider problems with inequality constraints for which the functions determining the objective and the constraints are all C^1. The convexity of this type of problem comes from the mere fact that the constraints are inequalities: $f_i(x) \leq 0, 1 \leq i \leq m$ can be reformulated as $f_i(x) + v_i = 0, v_i \geq 0, 1 \leq i \leq m$, that is, writing $F = (f_1, \ldots, f_m)^T$ and $v = (v_1, \ldots, v_m)^T$, we get the smooth constraint $F(x) + v = 0$ and the convex constraint v is contained in the first orthant of $\mathbb{R}^n$, a convex set.

Basic properties of convexity. A subset A of $\mathbb{R}^n$ is called convex if it contains along with each two points also the entire segment connecting these points: $x, y \in A, \alpha \in [0, 1] \Rightarrow (1 - \alpha)x + \alpha y \in A$. The intersection of a collection of convex sets is again a convex set. The convex hull of a subset $S \subseteq \mathbb{R}^n$ is the smallest convex set containing S; it can be described either as the set of all finite convex combinations of elements of S—$\sum_{k=1}^{N} \alpha_k x_k$ with $\alpha_k \in [0, 1]$, $k = 1, \ldots, N, \sum_{k=1}^{N} \alpha_k = 1, x_k \in S$, $k = 1, \ldots, N$—or as the intersection of all convex subsets containing S. A function $f : A \to \mathbb{R}$ with $A \subseteq \mathbb{R}^n$ is called convex if its epigraph $\{(x, \rho) | x \in A, \rho \in \mathbb{R}, \rho \geq f(x)\}$ is a convex subset of $\mathbb{R}^{n+1}$. This is equivalent to: A is a convex set and Jensen's inequality holds: $f((1-\alpha)x+\alpha y) \leq (1-\alpha)f(x)+\alpha f(y)$ for all $x, y \in A, \alpha \in [0, 1]$. One can prove that a convex function $f : A \to \mathbb{R}$ is continuous on the interior of its domain A. A convex function $f : A \to \mathbb{R}$ is called strictly convex if its graph does not contain any segments. This is equivalent to $f((1 - \alpha)x + \alpha y) < (1 - \alpha)f(x) + \alpha f(y)$ for

$x, y \in A, x \neq y, \alpha \in (0,1)$. For C^2-functions $f : \mathbb{R}^n \to \mathbb{R}$ one has the following criterion for convexity: the hessian matrix $f''(x)$, a symmetric matrix, is positive semi-definite for all $x \in \mathbb{R}^n$. One has the following *sufficient* condition for strict convexity: $f''(x)$ is positive definite for all $x \in \mathbb{R}^n$. If f is a quadratic function, $f(x) = x^T Ax + bx + c$ with A a symmetric $n \times n$-matrix, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$, then $f''(x) = 2A$ does not depend on x; then it can be verified whether f is convex (resp. strictly convex) by reducing the matrix A to diagonal form 'by symmetric sweeping' (carrying out repeatedly a suitable row operation and then the 'same' column operation) and checking whether all diagonal entries of the outcome are nonnegative (resp. positive).

Separation. A hyperplane H in $\mathbb{R}^n$ is a solution set of a linear equation $\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta$ for which $\alpha_i \neq 0$ for at least one index i. This hyperplane is said to separate two subsets S and T of $\mathbb{R}^n$ if each one lies on a different side of the hyperplane (both sets are allowed to have common points with the hyperplane); that is, either $s \in S, t \in T \Rightarrow \alpha_1 s_1 + \cdots + \alpha_n s_n \geq \beta, \alpha_1 t_1 + \cdots + \alpha_n t_n \leq \beta$ or $s \in S, t \in T \Rightarrow \alpha_1 s_1 + \cdots + \alpha_n s_n \leq \beta, \alpha_1 t_1 + \cdots + \alpha_n t_n \geq \beta$.

Separation theorem. *A nonempty closed convex set $A \subseteq \mathbb{R}^n$ and a point $p \in \mathbb{R}^n \backslash A$ can be separated by a hyperplane in $\mathbb{R}^n$.*

Sketch proof. Consider the distance problem for p and A: $|x - p| \to \min, x \in A$. This problem has a global solution; choose a point $\bar{a} \in A$; add the constraint $|x - p| \leq |\bar{a} - p|$ to the problem (this does not change the—possibly empty—solution space of the problem and it makes the admissible set bounded), and apply Weierstrass. This shows that the distance problem has a solution $\widehat{x}$. Then take the hyperplane H through q that is orthogonal to the line through $\widehat{x}$ and $\widehat{x}$. This hyperplane separates p and A, as a picture makes intuitively clear (make a picture!) and as can be readily verified by a calculation.

Remark. The conclusion of the theorem still holds if p is a boundary point of A, that is, if there are arbitrary close to p points from A as well as points that are not from A. To see this, one takes a sequence of points that are not from A, with limit p. Each point in the sequence can be separated from A by a hyperplane, by the separation theorem above. Consider the resulting sequence of hyperplanes. Take a convergent subsequence. Its limit has the required property: it separates p and A.

***The most general separation result.** We mention here the most general separation result, for use in other courses (Microeconomics). It is a criterion for the existence of a hyperplane that separates two convex sets A and B in $\mathbb{R}^n$: the existence of a point in $\mathbb{R}^n$ that is a relative interior point of A as well as of B. Relative interior point means interior point relative to the affine hull of the set in question; the affine hull of a set $S \subseteq \mathbb{R}^n$ is the affine space $S - S = \{s - t, s, t \in S\}$; this is called an affine space as it is of the form $p + L = \{p + l | l \in L\}$ for some $p \in V$ and some linear subspace L of $\mathbb{R}^n$. This general result is an immediate corollary of the special cases of a point and a convex set given above: separating A and B is equivalent to separating the origin from the convex set $A - B = \{a - b | a \in A, b \in B\}$: if $L : \mathbb{R}^n \to \mathbb{R}$ is linear and nonzero, then $L(a) \geq L(b) \, \forall a \in A \, \forall b \in B \Leftrightarrow L(a - b) \geq 0 \, \forall a \in B \, \forall b \in B$.

Subdifferential. The subdifferential of a convex function $f : \mathbb{R}^n \to \mathbb{R}$ in a point $\widehat{x}$ is defined to be

$\partial f(\widehat{x}) = \{v \in (\mathbb{R}^n)^T | f(\widehat{x}+h) - f(\widehat{x}) \geq vh \forall h \in \mathbb{R}^n\}$; here vh is the matrix product of the row vector v and the column vector h. Elements in the subdifferential are called subgradients. By the remark after the separation theorem, the subdifferential is always nonempty: the point $(\widehat{x}, f(\widehat{x}))$ is a boundary point of the epigraph of f, so $(\widehat{x}, f(\widehat{x}))$ and the epigraph of f can be separated by a hyperplane. This hyperplane can be viewed as the graph of a function that is affine (linear plus constant); its slope is then a subgradientb of f in $\widehat{x}$. Moreover, one can prove: the subdifferential consists of exactly one element iff f is differentiable in $\widehat{x}$; then the unique element of the subdifferential is the derivative $f'(\widehat{x})$.

Intuition subdifferential. What the derivative is for a differentiable function is the subdifferential for a convex function.

Convex optimization. An optimization problem is called convex if it is of the form $f(x) \to \min, x \in A$ with $A \subseteq \mathbb{R}^n$ a set and $f : A \to \mathbb{R}$ a convex function (and so in particular A is a convex set).

Proposition. *Each local minimum $\widehat{x}$ of a convex optimization problem $f(x) \to \min, x \in A$ is a global minimum.*

Proof. We argue by contradiction. Assume there is $x \in A$ with $f(x) < f(\widehat{x})$. Then the interval with endpoints $(\widehat{x}, f(\widehat{x}))$ and $(x, f(x))$ lies entirely below the horizontal plane of level $f(\widehat{x})$; on the other hand this lies in the epigraph of f. This implies that there are in this epigraph, arbitrary close to the point $(\widehat{x}, f(\widehat{x}))$, points of level lower than $f(\widehat{x})$. This contradicts the local minimality of $\widehat{x}$.

2.2 Convex Fermat

Convex variant of the theorem of Fermat. *Ingredients: $f : \mathbb{R}^n \to \mathbb{R}$, $\widehat{x} \in \mathbb{R}^n$. Assumption: f convex. Conclusion: a criterion for global minimality of $\widehat{x}$ for the problem $f(x) \to \min, x \in \mathbb{R}^n$ is that the subdifferential of f in $\widehat{x}$ contains the origin: $0_n \in \partial f(\widehat{x})$.*

Proof. Immediate consequence of the definition of the subdifferential.

Calculus for subdifferentials. The value of this result lies in its use combined with calculus rules. Here are the main ones: (1) basic examples: (i) $\partial f(x) = \{f'(x)\}$ if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable in $x \in \mathbb{R}^n$, if $N(x) = |x| = \sqrt{\sum_{k=1}^n x_k^2}$, then $\partial N(x) =$ the closed unit disk in $\mathbb{R}$ for $x = 0_n$, and it equals $\{\frac{x^T}{|x|}$ for all $x \neq 0^n$, (2) rules to calculate new subdifferentials from old ones if $f, g : \mathbb{R}^n \to \mathbb{R}$ are convex functions and $x \in \mathbb{R}^n$: (i) $\partial(f+g)(x) = \partial f(x) + \partial g(x)$ (the sum of two subsets A, B of $\mathbb{R}^n$ is defined by $A + B = \{a + b | a \in A, b \in B\}$), (ii) $\partial \max(f, g)(x)$ is the convex hull of $\partial f(x)$ and $\partial f(x)$ if $f(x) = g(x)$ and otherwise it equals $\partial f(x)$ if $f(x) > g(x)$ and $\partial g(x)$ if $g(x) > f(x)$.

Intuition calculus subdifferentials. Definitions are rarely used in calculations. This applies in particular to subdifferentials. In a point of differentiability, we just have to calculate the derivative.

In a point of non-differentiability, one has to know one basic example only: the subdifferential of the euclidean norm taken in the origin: this is the unit disk. Moreover, one has to look at the building pattern of the given convex function and for each building step one has to know the corresponding rule for the subdifferential. We only consider two building steps: adding convex functions and taking the maximum of two convex functions. The rules are given above. Carrying out these rules is a challenge, as these require the tasks of determining the sum $A+B$ of two given sets and the convex hull $\mathrm{co}(A,B)$ of two convex sets. You will be asked to carry out these rules for convex sets in the two-dimensional space, the plane, only; then you have to develop some 'drawing skilsl' for the operations $A+B$ and $\mathrm{co}(A,B)$.

Numerical illustration. Solve the problem $x^2+2y^2-3x-y+2\sqrt{x^2-4x+y^2+4} \to \min$.

1. Modeling and convexity. $f(x,y) = (x \quad y)\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} - 3x - y + 2|(x,y)-(2,0)| \to \min, (x,y) \in (\mathbb{R}^2)^T$. The function $f(x,y)$ is strictly convex, as its building pattern reveals: it is the sum of the strictly convex function $(x \quad y)\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$ and the two convex functions $-3x-y$ and $2|(x,y)-(2,0)|$.

2. Criterion. Convex Fermat: $0 \in \partial f(x,y)$. There is a unique point of non-differentiability, $(2,0)$; in this point, f has subdifferential the sum of the derivative $(x^2+2y^2-3x-y)' = (2x-3, 4y-1)$ taken in the point $(2,0)$—that is, $(1,-1)$—and the unit disk multiplied from the origin with the scalar 2. That is, we have to take the disk with center the origin and radius 2 and translate it over the vector $(1,-1)$. This gives the disk with center at $(1,-1)$ and radius 2.

3. Analysis. Let us try the point of non differentiability to begin with. It is a special point. Its subdifferential is a relatively large set so we have a fair chance that it contains the origin. All other points have a subdifferential consisting of one point only. It would be unlikely that in a randomly chosen such points the derivative is zero. Trying to find a point of differentiability that is a solution requires therefore the solution of the stationarity equations. These are two complicated equations in two unknowns. Let us hope that we are lucky with the investigation of the point of non-differentiability. Then we do not have to solve this system. Well, the criterion for optimality of the point of non-differentiability $(2,0)$ is: the origin has to lie inside the disk with center $(1,-1)$ and radius 2. The distance between $(1,-1)$ and $(0,0)$ is $\sqrt{2}$, which is indeed smaller than 2.

4. Conclusion. The problem has a unique solution $(2,0)$.

Convincing illustration. The facility location problem of Fermat-Weber. A facility has to be placed for which the average distance to three given points $a_i, i=1,2,3$—that do not lie on one straight line—is minimal. The usual textbook solution is by the theorem of Fermat; see for example B-T section 2.2.4. The cases that a Torricelli point does not exist are often omitted or require an awkward additional analysis. Below we use the convex Fermat theorem, following Vladimir Protassov, and this gives an efficient treatment that includes all cases.

1. Modeling and existence: $f(x) = \frac{1}{3}\sum_{i=1}^{3} |x - a_i| \to \min, x \in \mathbb{R}^2$, $f(x) \approx |x|^2 \to \infty$ for $|x| \to \infty$, so by Weierstrass a global solution $\widehat{x}$ exists. The function f is strictly convex.

2. Convex Fermat: $0 \in \partial f(x)$; this subdifferential equals $\frac{1}{3}\sum_{i=1}^{3} \frac{x-a_i}{|x-a_i|}$ if $x \neq a_1, a_2, a_3$ and it equals $\{u||u| \leq 1\} + \frac{a_i-a_j}{|a_i-a_j|} + \frac{a_i-a_k}{|a_i-a_k|}$ for $x = a_i$, where we have written $\{a_i, a_j, a_k\} = \{a_1, a_2, a_3\}$.

3. Analysis: in a point of non-differentiability $x = a_i$ we get (use: sum of two unit vectors has length ≤ 1 iff these make an angle $\geq 2\pi/3$) that the angle of the triangle with vertices a_1, a_2, a_3 at a_i is $\geq 2\pi/3$; in a point of differentiability we get (use: the sum of three unit vectors equals zero iff these make angles $2\pi/3$ with each other) that x is the Torricelli point of the triangle, the point that sees each pair of vertices of the triangle under the same angle—$2\pi/3$.

4. Conclusion: if one of the angles of the triangle is $\geq 2\pi/3$, then the unique optimal location is at the vertex of that angle; otherwise there exists a unique Torricelli point and this is the unique optimal location.

Intuition convex Fermat. If we want to minimize a function without constraints, then our first reflex might be to use the Fermat theorem. However, if the function is convex, it is better to use the convex variant of the Fermat theorem instead: then we get a *criterion* for *global* minimality. Having a criterion is of course the ideal situation. In particular, we get a criterion also for points of non-differentiability (warning: calculating the subdifferential at these points is a challenge). Moreover, then we have the additional advantage that we do not have to use the Weierstrass theorem, provided we can solve the stationarity equations explicitly. If we cannot solve these explicitly, but only get a characterization, such as in Fermat-Weber or in the oil pipeline for example, then we do have to use Weierstrass. So then we do not have this additional advantage.

2.3 Karush Kuhn Tucker

Intuition Karush Kuhn Tucker. In the lecture on smooth optimization, we have considered the problem with equality constraints; here the functions determining the objective and the constraints are all assumed to be differentiable. Now we consider the equivalent of this problem for convex optimization. It is the problem with inequality constraints; the functions determining the objective and the constraints are all assumed to be convex. Here we have again the ideal situation: a criterion for optimality. It is a variant of the Lagrange equations, called the Karush Kuhn Tucker conditions.

The Lagrange equations are that we have a stationary point of the Lagrange function. The KKT-conditions are that (1) we have a point of minimum of the Lagrange function, (2) all multipliers are nonnegative, and (3) the multipliers of the non-active constraints are even zero. This last requirement is usually formulated by means of complementarity conditions $\lambda_i f_i(x) = 0, i = 1, \ldots, m$ (check that this is the same). We will use a novel presentation of the KKT conditions in terms of subgradients. This has advantages, as we will see.

Necessary, sufficient and Slater conditions. We consider the convex optimization problem wityh inequality constraints $(P) : f_0(x) \to \min, f_i(x) \leq 0, 1 \leq i \leq m$, where $f_i : \mathbb{R}^n \to \mathbb{R}, 0 \leq i \leq m$, and $\widehat{x} \in \mathbb{R}^n$ with $f_i(\widehat{x}) \leq 0, 1 \leq i \leq m$. We assume that the functions $f_i, 0 \leq i \leq m$, are convex. We define three conditions on an admissible point $\widehat{x}$ of (P); we need these to formulate the KKT conditions.

- Necessary condition (Π). The convex hull of the subdifferentials in $\widehat{x}$ of the objective function f_0, and the functions $f_i, i \in I$, from the constraints that are active at $\widehat{x}$, that is, $I = \{i \in \{1, \ldots, m\} | f_i(\widehat{x}) = 0\}$ contains the origin: $0_n^T \in \mathrm{co}(\partial f_0(\widehat{x}), \partial f_i(\widehat{x}),\ i \in I)$.
- Sufficient condition (Π_0). Same as (Π) but with the additional requirement that $i = 0$ is genuinely taking part in a convex combination of points in the subdifferentials $\partial f_i(\widehat{x}),\ i \in I \cup \{0\}$ that equals the origin.
- Slater condition. There exists an admissible point $\bar{x}$ for which all inequality constraints hold strictly, $f_i(\bar{x}) < 0\ \forall i \in \{1, \ldots, m\}$. Such a point $\bar{x}$ is called a Slater point.

Theorem of Karush Kuhn Tucker (in Protassov form). *Consider the convex optimization problem with inequality constraints (P) and an admissible point $\widehat{x}$.*

1. *Necessary condition. Condition (Π) is a necessary condition for optimality of $\widehat{x}$.*
2. *Sufficient condition. Condition (Π_0) is a sufficient condition for optimality of $\widehat{x}$.*
3. *Criterion. If the Slater condition holds, then (Π) is a criterion for optimality of $\widehat{x}$.*

For the traditional form of the theorem of Karush Kuhn Tucker, see theorem 4.4 in [B-T].

Numerical illustration. Solve the problem $y^4 - y - 4x \to \min, \max\{x + 1, e^y\} + 2\sqrt{x^2 + y^2} \leq 3y + 2\sqrt{5}, \sqrt{x^2 + y^2 - 4x - 2y + 5} + x^2 - 2y \leq 2$.

1. Modeling and convexity. $f(x, y) = y^4 - y - 4x \to \min, (x, y) \in (\mathbb{R}^2)^T, g(x, y) = \max\{x+1, e^y\} + 2|(x, y)| - 3y - 2\sqrt{5} \leq 0, h(x, y) = |(x, y) - (2, 1)| + x^2 - 2y - 2 \leq 0$. This is seen to be a convex optimization problem and the objective function is strictly convex.

2. Criterion: the origin is contained in the convex hull of $\partial f(x,y)$, $\partial g(x,y)$ and $\partial h(x,y)$.

3. Analysis. Let us begin by investigating the point of non-differentiability $(2,1)$. At this point we have : (1) the subdifferential of f consists of the point $f'(2,1) = (-4,3)$, (2) the first constraint is active and the subdifferential of g consists of the point $g'(2,1) = (1,1) = 2\frac{(2,1)}{|(2,1)|} + (0,-3) = (1+\frac{4}{5}\sqrt{5}, -2+\frac{2}{5}\sqrt{5})$, (3) the second constraint is active,and the subdifferential of h is $D+(4,-2)$, where D denotes the unit disk. How to compute the convex hull of these subdifferentials? We can try to do this graphically by drawing a picture. However, is a picture reliable? Maybe not, but it gives us a valuable idea. We see that the origin lies more or less between the point $(-4,3)$, the subdifferential of f and the disk with center $(4,-2)$ and radius 1, which is the subdifferential of h. This suggests the following rigorous observation: the origin is the midpoint of the segment with endpoints $(-4,3)$ and $(4,-3)$,and the latter point lies in the the disk with center $(4,-2)$ and radius 1. This shows that the origin lies in the convex hull of the subdifferentials (and the subdifferential of the objective function f is genuinely taking part). That is, condition (Π_0) holds true and so $(2,1)$ is minimal.

4. Conclusion. The problem has a unique solution, $(\widehat{x}, \widehat{y}) = (2,1)$.

Convincing illustration. The shortest distance problem for an ellipse and a point outside the ellipse, serves to illustrate the advantage of KKT over Lagrange, see Problem 9.4.1 in [B-T]

*__Proof.__ We only prove the implication $\widehat{x} \in \mathrm{sol}(P) \Rightarrow (\Pi)$. We consider the family of problems $(P_y) : f_0(x) \to \min, f_i(x) \leq y_i, 1 \leq i \leq m$, where the parameter y runs over $\mathbb{R}^m$.

Let $S(y) \in \mathbb{R} \cup \{\pm\infty\}$ be the optimal value of (P_y), for each $y \in \mathbb{R}^m$. Then the function S is convex because of the convexity of the $f_i, 0 \leq i \leq m$. Apply the remark after the separation theorem to the point $(0, S(0))$ and the epigraph of S. This gives the existence of a hyperplane in $\mathbb{R}^{m+1}$ that contains the point $(0, S(0))$ and that has the epigraph entirely on one of its two sides. The side of the hyperplane that contains epi$S$ can be described by a linear inequality of the form $\lambda_0(y_0 - S(0)) + \sum_{i\in I} \lambda_i y_i \geq 0$ with $\lambda_i \geq 0\ \forall i \in I \cup \{0\}$ and $\sum_{i\in I} \lambda_i = 1$. Writing this out explicitly gives that the Lagrange function $\mathcal{L}(x) = \sum_{i\in I\cup\{0\}} \lambda_i f_i(x)$ is minimal in $\widehat{x}$ (to be precise: we have taken $\lambda_i = 0$ for all $i \in \{1,\ldots,m\}\backslash I$). This is equivalent, by the convex variant of the theorem of Fermat, to $0_n^T \in \partial\mathcal{L}(\widehat{x})$. Now we use repeatedly the calculus rule $\partial(f+g) = \partial f + \partial g$: this gives $0_n^T \in \sum_{i\in I\cup\{0\}} \lambda_i \partial f_i(\widehat{x})$, as required.

This proof gives the following shadow-price interpretation for the multipliers in the KKT conditions.

Theorem (shadow-price interpretation for multipliers of convex problems). *Consider the ingredients and assumptions of the theorem of Karush Kuhn Tucker. Assume that the KKT conditions hold for the point $\widehat{x}$ with choice of multipliers $\bar{\lambda} = (\lambda_0, \ldots, \lambda_m)$ with $\lambda_0 = 1$. Let $S(y)$ be the value of the problem $(P_y) : f_0(x) \to \min, f_i(x) \leq y_i, 1 \leq i \leq m$. Then $(-\lambda_1, \ldots, -\lambda_m) \in \partial S(0)$.*

The proof also gives the following form of the theorem of Karush Kuhn Tucker.

Theorem (duality theory). *Consider the ingredients and assumptions of the theorem of Karush*

Kuhn Tucker. Assume that the KKT conditions hold for the point $\widehat{x}$ with choice of multipliers $\bar{\lambda} = (\lambda_0, \ldots, \lambda_m)$ with $\lambda_0 = 1$. Let $S(y)$ be the value of the problem $(P_y) : f_0(x) \to \min, f_i(x) \leq y_i, 1 \leq i \leq m$. The dual problem is defined to be the problem to find $\eta \in (\mathbb{R}^m)^T$ such that the highest affine (= linear plus constant) function with slope η that lies under the epigraph of S has maximal intercept (= value at 0_m). Conclusion: the dual problem has solution $(\lambda_1, \ldots, \lambda_m)$ and its value equals the value of the given problem $f_0(x) \to \min, f_i(x) \leq 0, 1 \leq i \leq m$.

Intuition duality theory. Each convex set, convex function and convex optimization problem has two equivalent descriptions, one from the inside (called the primal description) and one from the outside (the dual description). For a closed convex set, the primal description is as the set of all points it contains, the dual description is as the intersection of all closed halfspaces that contain it. For a convex function f that is closed (that is, the epigraph is a closed set) epigraph, the primal description is as f, the dual description is by means of its conjugate function f^* which is defined by $f^*(y) = \sup_{x\in\mathbb{R}}(\langle x, y\rangle - f(x))$. This is again a closed convex function and taking its conjugate brings us back to the given function—$f^{**} = f$. For a convex optimization problem, the primal and dual description are given in the statement of the theorem. It is a remarkable phenomenon that we have two optimization problems, over completely different spaces, one minimization and the other maximization, which have the same optimal value. In applications, one of the two is a problem for engineers—choose optimal production levels—the other one is a problem for economists—choosing optimal prices.

2.4 Fritz John

Theorem of Fritz John (in Protassov form). *Ingredients: $f_i : \mathbb{R}^n \to \mathbb{R}$, $0 \leq i \leq m$, and $\widehat{x} \in \mathbb{R}^n$ with $f_i(\widehat{x}) \leq 0$, $1 \leq i \leq m$. Assumptions: the functions f_i, $0 \leq i \leq m$, are C^1-functions. Problem $(P) : f_0(x) \to \min, f_i(x) \leq 0, 1 \leq i \leq m$. Definition condition (Π) on $\widehat{x}$: the convex hull of the derivatives in $\widehat{x}$ of the f_i, $i \in I \cup \{0\}$, where $I = \{i \in \{1, \ldots, m\} | f_i(\widehat{x}) = 0\}$, contains the origin. Conclusion: $\widehat{x} \in \mathrm{sol}(P) \Rightarrow (\Pi)$.*

For the traditional form of the theorem of Fritz John, see theorem 4.13 in [B-T].

Numerical illustration. Find the smallest positive constant D such that $\sqrt{x_1^2 + x_2^2} \leq D \sqrt[4]{x_1^4 + x_2^4}$.

1. Modeling and existence. $f(x) = -x_1^2 - x_2^2 \to \min, g(x) = x_1^4 + x_2^4 - 1 \leq 0$; a solution exists by Weierstrass as the objective function is continuous and the admissible set is nonempty, closed and bounded (the latter because $(x_1^2)^2 + (x_2^2)^2 \leq 1 \Rightarrow -1 \leq x_1, x_2 \leq 1$.

2. Necessary conditions. Fritz John: either $(0, 0) = f'(x) = (-2x_1, -2x_2)$ or $(0, 0)$ is contained in the convex hull of $f'(x) = (-2x_1, -2x_2)$ and $g'(x) = (4x_1^3, 4x_2^3)$.

3. Analysis. The Fritz John conditions give $x = 0$ or $(-2x_1, -2x_2) = \alpha(4x_1^3, 4x_2^3)$ for some $\alpha > 0$, so $x_1^4 = x_2^4$ and so $x_2 = \pm x_1$ so, by the constraint $x = (\pm\sqrt[4]{1/2}, \pm\sqrt[4]{1/2})$. Comparing f-values shows that the four points $x = (\pm\sqrt[4]{1/2}, \pm\sqrt[4]{1/2})$ are optimal. The value of the problem is seen to be $\sqrt[4]{2}$

4. Conclusion. There are four solutions $\widehat{x} = (\pm\sqrt[4]{1/2}, \pm\sqrt[4]{1/2})$; this gives that the required value is $D = \sqrt[4]{2}$.

****Convincing illustration off KKT.** We give the proof by Vladimir Protassov of the theorem of Minkowski on convex polyhedra. Let a polyhedron in $\mathbb{R}^d$ have k $d-1$-dimensional boundaries, we denote by S_i the $d-1$-dimensional volume of the i-th boundary, and by n_i the unit vector that is normal to this boundary in outward direction. We use the traditional form off Fritz John.

Theorem (Minkowski). *For each convex polyhedron $\sum_{i=1}^k S_i n_i = 0$. Conversely, for arbitrary positive numbers $a_1, \dots, a_k$ such that $\sum_{i=1}^k a_i n_i = 0$ there exists a unique convex polyhedron with outward normals n_i for which $S_i = a_i, i = 1, \dots, n$.*

This theorem consists of three statements: the necessary condition $\sum_{i=1}^k S_i n_i = 0$ for the existence of a polyhedron, its sufficiency and the uniqueness of the polyhedron. We prove all except the uniqueness.

Proof of the necessity. Let G be a given polyhedron. For an arbitrary point $x \in G$ we denote by $h_i(x)$ the distance from this point to the i-th boundary plane. As the volume of the pyramid with top x and with base the i-th boundary, is equal to $\frac{1}{d} h_i S_i$, we get $\sum_{i=1}^k h_i(x) S_i = d\mathrm{vol}(G)$. Differentiating both side of the given identity, we obtain $\sum_{i=1}^k n_i S_i = 0$.

Proof of the sufficiency. For any choice of nonnegative numbers $h = (h_1, \dots, h_k)$ we consider the polyhedron $G(h) = \{x \in \mathbb{R}^d | (n_i, x) \le h_i, i = 1, \dots, k\}$ and we denote by $V(h)$ its volume. We observe that $0 \in G(h)$ for all h. We consider the problem

$$-V(h) \to \min, \sum_{i=1}^k a_i h_i = 1, -h_i \le 0,\ i = 1, \dots, k. \tag{1}$$

In other words, among all polyhedra that are bounded by hyperplanes that are perpendicular to vectors n_i and having distance from the origin, such that $\sum_{i=1}^k a_i h_i = 1$, we search for the polyhedron with maximal volume.

The solution of problem (1) exists, as the set of admissible vectors is compact. We denote a solution by $\widehat{h}$. By the Lagrange multiplier rule, we have $\mathcal{L}_h(\lambda, \widehat{h}) = 0$, where

$$\mathcal{L} = -\lambda_0 V(h) - \sum_{i=1}^k \lambda_i h_i + \lambda_{k+1}(\sum_{i=1}^k a_i h_i - 1)$$

is the Lagrange function. Moreover, $\lambda_i \ge 0$ and $\lambda_i \widehat{h}_i = 0, i = 1, \dots, k$.

Differentiating with respect to h_i, we obtain

$$-\lambda_0 V_{h_i} - \lambda_i + \lambda_{k+1} a_i = 0,\ i = 1, \dots, k.$$

If for some j the hyperplane $\{x \in \mathbb{R}^d | (n_j, x) = \widehat{h}_i\}$ does not intersect the polyhedron $G(\widehat{h})$ in a $d-1$-dimensional boundary, that is, if it is "superfluous", then $V_{h_j} = 0$, hence $\lambda_j = \lambda_{k+1} a_j$. This is also true in the case $\lambda_0 = 0$. If $\lambda_j = 0$, then $\lambda_{k+1} = 0$, and so $\lambda_i = -\lambda_0 V_{h_j}$ for all i. If $\lambda_0 = 0$, then we obtain a contradiction to Neron, and if $\lambda_0 = 1$, then $\lambda_i = -V_{h_i}$ for all i. If the i-th phyperplane intersects the polyhedron in a $d-1$-dimensional boundary (and such, of course, will exist), then $V_{h_i} = S_i > 0$, and so $\lambda_i < 0$, which contradicts the condition of nonnegativity. So $\lambda_0 = 1$, and all k hyperplanes form a boundary, that is $V_{h_i} = S_i > 0$ for all i. Then

$$-S_i - \lambda_i + \lambda_{k+1} a_i = 0, \quad i = 1, \dots, k.$$

Multiplying the given equality by n_i summing over all i, and taking into consideration that $\sum_{i=1}^k S_i n_i = \sum_{i=1}^k a_i n_i = 0$, we obtain $\sum_{i=1}^k \lambda_i n_i = 0$. We denote by I the set of indices i, for which $\lambda_i > 0$. Then $\sum_{i\in I} \lambda_i n_i = 0$, and by virtue of the complementary slackness condition, $\widehat{h}_i = 0$ for all $i \in I$. On the other hand, the polyhedron $G(\widehat{h})$ has at least one interior point z for which $(n_i, z) < h_i$, $i = 1, \ldots, k$. Multiplying by λ_i and summing over all $i \in I$, we obtain $0 < \sum_{i\in I} \lambda_i \widehat{h}_i = 0$. The given contradiction means that I is empty, that is $\lambda_i = 0$, and so $S_i = \lambda_{k+1} a_i$ for all $i = 1, \ldots, k$. Therefore, for the polyhedron $G(\widehat{h})$ the $d-1$-dimensional volume of the i-th boundary is proportional to a_i. Carrying out a homothety with appropriate coefficient, we obtain a polyhedron for which the $d-1$-dimensional volume of the boundaries is equal to these numbers.

2.5 Problems on the lecture notes:

1. Prove that the collection of finite convex combinations of elements of some set $S \subseteq \mathbb{R}^n$ is a convex set.

2. Prove that the definition of convex functions given in the lecture notes is equivalent to the usual one, given by Jensen's inequality $f((1-\alpha)x + \alpha y) \le (1-\alpha)f(x) + \alpha f(y)$ for all $x, y \in \mathbb{R}^n$ and all $\alpha \in [0, 1]$.

3. Show by examples that a convex function need not be continuous on the boundary points of its domain. A rich and simple class of examples is found by considering the functions f that are defined on the closed unit disk $\{x \in \mathbb{R}^2 | x_1^2 + x_2^2 \le 1\}$ that take value zero on the interior $\{x \in \mathbb{R}^2 | x_1^2 + x_2^2 < 1\}$. Such a function is convex if and only if it takes nonnegative values at all points of the boundary $\{x \in \mathbb{R}^2 | x_1^2 + x_2^2 = 1\}$. Prove this. That is, each choice of an arbitrary nonnegative valued function on the circle gives a continuous function. Give an explicit example of such a function on the boundary that makes the convex function non-continuous in each point of the boundary. [Hint: let the function on the boundary take two values depending on whether x_1 is a rational number or not.]

   Give a criterion for convexity for functions $f : D \to \mathbb{R}$ where $D = \{(x, y) | x^2 + y^2 \le 1\}$, the closed unit disk, for which $x^2 + y^2 < 1 \Rightarrow f(x, y) = 0$. Use this to show that there are convex functions $D \to \mathbb{R}$ that are not continuous.

4. Show by symmetric sweeping that the 3×3-matrix $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 8 \end{pmatrix}$ is positive definite.

5. Prove that for a closed convex set $A \subseteq \mathbb{R}^n$ and a point $p \in \mathbb{R}^n \setminus A$, and with $\widehat{x}$ the point in A that is closest to p, the hyperplane through the point $\widehat{x}$ that is orthogonal to the line through p and $\widehat{x}$ separates p and A. [Hint. By the minimality property of $\widehat{x}$ we have $|\alpha a + (1-\alpha)\widehat{x} - p| \ge |\widehat{x} - p| \,\forall \alpha \in [0, 1]$; take squares of both sides, use $|v|^2 = \langle v, v\rangle$, simplify, divide by α and take the limit $\alpha \downarrow 0$. Show that the inequality that is the outcome gives the required result.]

6. Determine the subdifferential of the euclidean norm function on $\mathbb{R}^2$ given by $x \mapsto |x| = \sqrt{x_1^2 + x_2^2}$ at the origin. [Hint: (c_1, c_2) lies in the subdifferential if $\sqrt{x_1^2 + x_2^2} \geq c_1x_1 + c_2x_2$ for all (x_1, x_2). Show that this is equivalent to $x_1^2 + x_2^2 \geq (c_1x_1 + c_2x_2)^2$ for all (x_1, x_2). Show that this in turn is equivalent to: A is positive semi definite, where A is the symmetric 2×2-matrix for which $x_1^2 + x_2^2 - (c_1x_1 + c_2x_2)^2 = x^T Ax$. Then use the minor criterion for positive semi definiteness.]

7. *Use the convex Fermat theorem to solve the more general Fermat-Weber location facility problem that the three clients are not equally important but have each a given weight (exercise 2.6.7).

8. Solve the problem $x^2 + 3y^2 - x - y + 2\sqrt{x^2 - 2x + y^2 + 1} \to \min$.

9. Solve the problem $x^2 - 2y \to \min$,
$\max\{3x^2, e^y + 2\} + \sqrt{x^2 + y^2 - 2x + 1} \leq 6x + \sqrt{5}$,
$\sqrt{x^2 + y^2 - 4x - 4y + 8} - 2x + 2y \leq 0$.

10. *In the proof of the theorem of KKT it is claimed that the convexity of the functions $f_i, 0 \leq i \leq m$, implies the convexity of the value function S. Prove this. [Hint. Show that the strict epigraph of S, $\{(y, \rho) | y \in \mathbb{R}, \rho \in \mathbb{R}, \rho > S(y)\}$ is convex and then you can use without proof the easy fact that this implies the convexity of S.] .

2.6 Your tasks between lecture 2 and lecture 3.

You are expected to do as follows.

- In groups of two you have to solve the ten problems on the lecture notes given above as well as the following five optimization problems in the following style (except for the first two numerical problems: here you should immediately carry out the four step method)—everything in your own words: begin by formulating the problem in words, then carry out the four step method and finally state the economic insight that the analysis gives (if any); you have to hand in your work on Tuesday 24-9-2013. Present complete solutions (not just some calculations). The problems are from [B-T]:

    1. D.4.2
    2. D4.5
    3. 1.6.6
    4. 1.6.14
    5. 1.6.33 [hint: use the convex Fermat]

- Study these lecture notes carefully (if something is not clear to you, look it up in [B-T]).

- Study carefully in [B-T] the royal roads as described on p.204 (ch 4), p.262 (ch5).
- Study carefully in [B-T] the following illustrations of the four step method: 1.4.4, 2.4.2, 2.4.3, 2.4.5.
- Look through the section in ch4 of [B-T] with applications.
- Study the texts about shadow price interpretations of Lagrange multipliers p. 244, 245.
- Study the symmetric sweeping algorithm for determining the definiteness of a symmetric matrix and try to find two ways to use it in the four step method.
- Study how the methods of optimization are applied to economic problems in chapter 8 (without worrying about the technical details).
- Look up in the Appendices what you need (for example about convex functions).
- Try to look through the chapters $4, 5$ and 8 in [B-T].
- Try to make the exercises from [B-T] without looking at the solutions at the back of the book; in any case the solutions you hand in should be your own texts.
- *Give a precise proof of the envelope theorem for $m = n = 1$; in particular, prove that $x(y)$ is differentiable, using the implicit function theorem (look this theorem up in your Analysis textbook for Math I, for example).

3 Lecture on dynamic optimization I: the Calculus of Variations (Euler, transversality, Euler-Lagrange), Optimal Control (Pontryagin's maximum principle)

3.1 Preliminary remarks on dynamic optimization.

A dynamic optimization problem is an optimization problem where the variable that has to be chosen in an optimal way is a time path of a quantity. Here time can be modeled discretely or continuously; for both cases there are three techniques available: the Calculus of Variations, Optimal Control, Dynamic Programming. The Calculus of Variations is a special case and a predecessor of Optimal Control. It is still used: often you do not need the full power of Optimal Control and then it is more convenient to use the Calculus of Variations. Moreover, it is a good idea to learn the Calculus of Variations as a preparation for Optimal Control. The Calculus of Variations and Optimal Control give necessary conditions (just like Fermat, Lagrange, KKT, FJ); Dynamic Programming gives sufficient conditions.

In your later TI courses you will see that for problems that are modeled in continuous time, the Calculus of Variations and Optimal Control are used; for problems that are modeled in discrete time Dynamic Programming is used. This is a choice; other choices are possible. As life does not end at the TI, we will give in this course also some attention, though much less, to other choices. The TI choice makes sense. In the first place, for continuous time models, the Calculus of Variations and Optimal Control lead to differential equations; Dynamic Programming leads to partial differential equations. Whereas we have a fighting chance to solve a differential equation, solving a given partial differential equation is much harder and is beyond the scope of this course. In the second place, for discrete time models, Dynamic Programming leads to a rich and often practical recursive method, whereas the Calculus of Variations and Optimal Control are essentially just Fermat, Lagrange, KKT, FJ in disguise.

Each of the three methods has one tool: (1) the Euler equation (and variants such as transversality conditions and the Euler-Lagrange equation) for the Calculus of Variations, (2) Pontryagin's Maximum Principle for Optimal Control, (3) the Bellman equation, also called Hamilton-Jacobi-Bellman equation, for Dynamic Programming.

All this might look intimidating at first glance, but this wrong impression can be removed from our mind when we look at things in the general context of optimization theory. The situation is very clear: there are two results only that concern us here: a necessary condition and a sufficient condition for a very general type of problem. All problems that concern us in this course are special cases of this general type. All the necessary conditions that we use are special cases of the general necessary condition. So Fermat, Lagrange, KKT, FJ, Euler, Euler-Lagrange, transversality, Pontryagin's Maximum Principle are all special cases of one general result (see [B-T] Chapter 10 for details). On the other hand, the first and second order necessary conditions given in lecture 1 and the Bellman equation are both special cases of one general sufficient condition. Treatment of the necessary and sufficient condition for this general type of problem is beyond the scope of this course; however, it is hopefully a comforting thought that the seemingly bewildering collection of conditions that confront us (Fermat, Lagrange, KKT, FJ, Euler, Euler-Lagrange, Pontryagin's Maximum Principle, Bellman ...) are just the result of what you get by mechanically working out two general results in order to get practically usable formulas for special problem types of interest. It is also possible to derive existence results from the general compactness principle that was referred to in the first lecture, but this is beyond the scope of this course. **Recommendation: conclude each analysis of a problem with the methods of the Calculus of Variations or Pontryagin's Maximum Principle by using the phrase 'If there exists a solution, then it is ...'.**

In this lecture we will discuss the first two methods, the Calculus of Variations and Optimal Controol, for continuous time models (and briefly for discrete time models).

3.2 The Calculus of Variations: continuous time

We consider the problem

$$J(x(\cdot),u(\cdot)) = \int_{t_0}^{t_1} f(t,x(t),u(t))dt + \kappa(x(t_1)) \to \min, \dot{x} = \varphi(t,x(t),u(t)), x(t_0) = x_0, (x(t_1) = x_1).$$

That is, we consider two variants: one with final value $x(t_1)$ fixed and one with final value of x free. Here $t_0, t_1 \in \mathbb{R}, t_0 \leq t_1; f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}, \varphi : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$, and $\kappa : \mathbb{R}^n \to \mathbb{R}$ are C^1-functions, $x_0 \in \mathbb{R}^n$, $(x_1 \in \mathbb{R}^n)$. We let $x(\cdot)$ run over all C^1-functions $[t_0, t_1] \to \mathbb{R}^n$ and $u(\cdot)$ over the continuous functions $[t_0, t_1] \to \mathbb{R}^r$. This problem is called the Lagrange problem.

Intuition Lagrange problem. This problem can be viewed as follows: if we choose a time path $u(\cdot)$ for the decision variable—that is we make decision $u(t)$ at time t for each $t \in [t_0, t_1]$—then the Cauchy problem $\dot{x} = \varphi(t,x(t),u(t)), x(t_0) = x_0$ has, under mild assumptions, a unique solution $x(\cdot)$ by the main theorem of differential equations—the Picard theorem. So the objective function $J(x(\cdot), u(\cdot))$ depends really on one choice, that of the decision function $u(\cdot)$. The Cauchy problem can be viewed as modeling some law of growth for the function $x(\cdot)$ under the influence of human decisions, represented by the function $u(\cdot)$.

Theorem (Euler-Lagrange equations and transversality condition). *Consider the Lagrange problem given above. Define the Hamiltonian $H = H(t, x, u, p, \lambda_0) = p\varphi - \lambda_0 f$. Conclusion: for each* *solution $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ there is a number $\widehat{\lambda}_0 \geq 0$ and a C^1-function $\widehat{p}(\cdot) : [t_0, t_1] \to \mathbb{R}^n$, not both zero, such that (again in shorthand notation) the Euler-Lagrange equations $\dot{\widehat{x}} = \widehat{H}_p, \dot{\widehat{p}} = -\widehat{H}_x, \widehat{H}_u = 0$, hold, as well as the transversality condition $\widehat{p}(t_1) = -\lambda_0 \widehat{\kappa}'(t_1)$ (the transversality condition only if the final value of x is free).*

Here we have used shorthand notation: for example $\widehat{H}_p$ stands for $\frac{\partial H}{\partial p}(t, \widehat{x}(t), \widehat{u}(t), \widehat{p}(t), \widehat{\lambda}_0)$.

On the proof. For simplicity, take $m = n = 1$ and $x_0 = 0$; we consider the case with endstate $x(t_1)$ free. Let us be bold and just apply the multiplier rule. We define the Lagrange function to be $\mathcal{L} = \lambda_0(\int_{t_0}^{t_1} f(t, x(t), u(t))dt + \kappa(x(t_1))) + \langle p, \dot{x} - \varphi(t, x, u)\rangle$. Here λ_0 is a number and $p : [t_0, t_1] \to \mathbb{R}^n$ is a C^1-function,; the inner-product notation means: take the product and integrate from t_0 to t_1: that is, $\langle p, \dot{x} - \varphi(t, x, u)\rangle = \int_{t_0}^{t_1} p(t)(\dot{x}(t) - \varphi(t, x(t), u(t)))dt$. Then we show that the stationarity equations of the Lagrange function with respect to the variable functions $x(\cdot)$ and $u(\cdot)$ are the second, third and last equation in the statement of the theorem (the first equation is just the constraint $\dot{\widehat{x}} = \widehat{\varphi}$ so it needs no proof). To be more precise, it suffices to establish the following formula:

$$\mathcal{L}(\widehat{x} + h, \widehat{u} + k) - \mathcal{L}(\widehat{x}, \widehat{u}) \approx \langle -\dot{\widehat{p}} - \widehat{H}_x, h\rangle + \langle -\widehat{H}_u, k\rangle + (\widehat{p}(t_1) + \widehat{\lambda}_0 \kappa'(\widehat{x}(t_1)))h(t_1)$$

for all C^1-functions h on $[t_0, t_1]$ with $h(t_0) = 0$ (this is needed to ensure that $\widehat{x} + h$ is admissible as well) and such that h and its derivative take on values on $[t_0, t_1]$ that are ≈ 0, and all continuous functions k on $[t_0, t_1]$ that take on values ≈ 0 on $[t_0, t_1]$.

The following calculation shows this:

$$\mathcal{L}(\widehat{x} + h, \widehat{u} + k) - \mathcal{L}(\widehat{x}, \widehat{u}) \approx$$

(by taking first order approximations under the integral sign)

$$\int_{t_0}^{t_1} [\widehat{\lambda}_0 \widehat{f}_x(t)h(t) + \widehat{p}(t)(\dot{h(t)} - \widehat{\varphi}_x(t)h(t)) + (\widehat{\lambda}_0 \widehat{f}_u(t) - \widehat{p}(t)\widehat{\varphi}_u(t))k(t)]dt + +\widehat{\lambda}_0 \kappa'(\widehat{x}(t_1))h(t_1) =$$

(by partial integration and using $h(t_0) = 0$)

$$\int_{t_0}^{t_1} [(\widehat{\lambda}_0 \widehat{f}_x(t) - \dot{\widehat{p}}(t) - \widehat{\varphi}_x(t))h(t) + (\widehat{\lambda}_0 \widehat{f}_u(t) - \widehat{p}(t)\widehat{\varphi}_u(t))k(t)]dt + \widehat{p}(t_1)h(t_1) + \widehat{\lambda}_0 \kappa'(\widehat{x}(t_1))h(t_1) =$$

(using the notation of the Hamiltonian and of the inner product)

$$\langle -\dot{\widehat{p}} - \widehat{H}_x, h\rangle + \langle -\widehat{H}_u, k\rangle + (\widehat{p}(t_1) + \widehat{\lambda}_0 \kappa'(\widehat{x}(t_1)))h(t_1),$$

as required.

Numerical illustration. Speed up the growth of a plant by artificial lighting with minimal cost in order to reach a required length at a given time. We model the law of growth by $\dot{x} = 1$ and the influence of lighting u by $\dot{x} = 1 + u$, the requirement by $x(1) = 2$, and the cost of lighting $\int_0^1 u^2 dt$.

1. Modeling. $J(x(\cdot), u(\cdot)) = \int_0^1 u^2 dt \to \min, \dot{x} = 1 + u, x(0) = 0, x(1) = 2$.

2. Necessary conditions. Hamiltonian: $H = p\varphi - \lambda_0 f = p(1+u) - \lambda_0 u^2$. Euler Lagrange: $\dot{x} = H_p = 1 + u, \dot{p} = -H_x = 0, 0 = H_u = p - 2\lambda_0 u$.

3. Analysis. Exclusion bad case: $\lambda_0 = 0 \Rightarrow p \equiv 0$ (using $H_u = 0$), contradiction. Now take $\lambda_0 = 1$.

   $\dot{p} = 0 \Rightarrow p \equiv c$, a constant; so $u = c/2$; so $\dot{x} = 1 + c/2$. It follows that $x(t) = At + B$ for suitable constants A, B. The boundary conditions $x(0) = 0, x(1) = 2$ give $A = 2, B = 0$

4. Conclusion. If there exists a solution, then it is $\widehat{x}(t) = 2t$.

Convincing illustration. Growth theory problems (see for example section 12.3.1 from [B-T]).

Intuition Euler-Lagrange and transversality. For each type of extremal problems one can get necessary conditions by the same principle—the method of Lagrange multipliers. Here the constraint is an equality of functions $\dot{x} = \varphi(t, x, u)$; this holds for each t, so for each t one needs a multiplier $p(t)$. In other words the multiplier is a function $t \mapsto p(t)$. To form the Lagrange function, one should not take the sum of products of multipliers and left hand sides of equality constraints $p(t)(dotx(t) - f(t, x(t), u(t))$ over all $t \in [t_0, t_1]$, but the integral over $[t_0, t_1]$. This gives the Lagrange function $\mathcal{L}$ given above. Then the Lagrange principle then prescribes to write the stationarity conditions of the problem to minimize the Lagrange function, where $x(\cdot)$ and $u(\cdot)$ are free and not bound by the constraint $\dot{x} = \varphi(t, x, u)$. To this end we use the approximation definition of the derivative: a calculation shows conditions are that $\mathcal{L}(\widehat{x} + h, \widehat{u} + k) - \mathcal{L}(\widehat{x}, \widehat{u})$ can be split up as the sum of $\langle -\dot{\widehat{p}} - \widehat{H}_x, h\rangle + \langle -\widehat{H}_u, k\rangle + (\widehat{\lambda}_0 \widehat{\kappa}'(t_1) + \widehat{p}(t_1))h(t_1)$, which is linear in $(h(\cdot), k(\cdot))$ and a term that is negligible in $(h(\cdot), k(\cdot))$. Therefore, the stationary conditions are $\widehat{p} - \widehat{H}_x = 0, -\widehat{H}_u = 0, \widehat{p}(t_1) = -\widehat{\lambda}_0 \widehat{\kappa}'(t_1)$, to which one can add $\dot{\widehat{x}} = \widehat{H}_p$, which is just the sonstraint $\dot{\widehat{x}} = \widehat{\varphi}$ in disguise. Thus we get the Euler equations and the transversality conditions.

Often one can eliminate u from the constraint $\dot{x} = \varphi(t, x, u)$; this gives a problem of the type

$$I(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt + \kappa(x(t_1)) \to \min, x(t_0) = x_0, (x(t_1) = x_1).$$

Then it is usual to formulate the necessary conditions without using the Hamiltonian. Here are the results for both case (endstate $x(t_1)$ prescribed and free)

Theorem (Euler equation). *Ingredients:* $t_0, t_1 \in \mathbb{R}, t_0 \le t_1, x_0, x_1 \in \mathbb{R}^n, L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. *Assumption:* L *is* C^2. *Problem:* $I(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \to \min, x(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n), x(t_0) = x_0, x(t_1) = x_1$. *Conclusion: a solution* $\widehat{x}(\cdot)$ *of the problem satisfies the Euler equation* $\widehat{L}_x = \frac{d}{dt}\widehat{L}_{\dot{x}}$ *(or written in full:* $L_x(t, \widehat{x}(t), \dot{\widehat{x}}(t)) = \frac{d}{dt} L_{\dot{x}}(t, \widehat{x}(t), \dot{\widehat{x}}(t)) \, \forall t \in [t_0, t_1]$*).*

Here we give a simple proof, which is due to Vladimir Tikhomirov. The usual proof, that is given in all textbooks, requires the relatively technical lemma of du Bois-Reymond.

Proof. For simplicity we assume $x_0 = x_1 = 0$. Choose a C^1-function $h(\cdot) : [t_0, t_1] \to \mathbb{R}^n$ with $h(t_0) = h(t_1) = 0$ (we call this an 'allowable $h(\cdot)$'). Then $\widehat{x}(\cdot) + \alpha h(\cdot)$ is admissible $\forall \alpha \in \mathbb{R}$ and $I(\widehat{x}(\cdot) + \alpha h(\cdot))$ is as a function of α minimal in $\alpha = 0$, by the minimality property of $\widehat{x}(\cdot)$. Apply Fermat: put the derivative equal to zero. This gives the equation $\int_{t_0}^{t_1}(\langle \widehat{L}_{\dot{x}}, \dot{h}\rangle + \langle \widehat{L}_x, h\rangle)dt = 0$, where $\langle a, b\rangle = \sum_{i=1}^n a_i b_i$. Partial integration gives $\int_{t_0}^{t_1}\langle \widehat{L}_{\dot{x}}(t) + \int_t^{t_1} \widehat{L}_x(\tau)d\tau, \dot{h}(t)\rangle dt = 0$. This can be written as

$$\int_{t_0}^{t_1} \langle \widehat{L}_{\dot{x}}(t) + \int_t^{t_1} \widehat{L}_x(\tau)d\tau - c, \dot{h}(t)\rangle dt = 0 \qquad (*)$$

for any $c \in \mathbb{R}$, as $h(t_0) = h(t_1)$ and so $\int_{t_0}^{t_1} \dot{h}dt = 0$. Now we choose $h(\cdot)$ and c such that $\dot{h}(t) = \widehat{f}_{\dot{x}}(t) + \int_t^{t_1} \widehat{L}_x(\tau)d\tau - c, h(t_0) = 0$ and $\int_{t_0}^{t_1} \dot{h}(t)dt = 0$. Then $h(t_1) = 0$, as $0 = \int_{t_0}^{t_1} \dot{h}dt = h(t_1) - h(t_0)$ and $h(t_0) = 0$, and so this $h(\cdot)$ is allowable. Substitution of this $h(\cdot)$ in $(*)$ gives $\int_{t_0}^{t_1}\langle g(t), g(t)\rangle dt = 0$ and so $g \equiv 0$, where $g(t) = \widehat{L}_{\dot{x}}(t) + \int_t^{t_1} \widehat{L}_x(\tau)d\tau - c$. It follows that $\widehat{L}_{\dot{x}}$ is continuously differentiable. Differentiation gives the Euler equation.

Convincing illustration. The brachistochrone problem (see [K-S]).

Theorem (transversality condition for free endstate). *Ingredients:* $t_0, t_1 \in \mathbb{R}, t_0 \leq t_1, x_0 \in \mathbb{R}^n, L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \kappa : \mathbb{R}^n \to \mathbb{R}$. *Assumption:* L *is* C^2 *and* κ *is* C^1. *Problem:* $I(x(\cdot)) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt + \kappa(x(t_1)) \to \min, x(\cdot) \in C^1([t_0, t_1], \mathbb{R}^n), x(t_0) = x_0$. *Conclusion: a solution* $\widehat{x}(\cdot)$ *of the problem satisfies the Euler equation* $\widehat{L}_x = \frac{d}{dt}\widehat{L}_{\dot{x}}$ *and the transversality condition* $\widehat{L}_{\dot{x}}(t_1) = -\widehat{\kappa}'(t_1)$.

Proof. Similar to the proof above.

Shadow-price interpretations of the function $\widehat{p}(\cdot)$ and the Hamiltonian function. We emphasize the dependence of the optimal value of the Lagrange problem on the pair (t_0, x_0) by denoting it $V(t_0, x_0)$. Then, if this function V is C^1, and $\widehat{\lambda}_0 = 1$ (for all pairs (t_0, x_0)), then we have $V_x(t_0, x_0) = -\widehat{p}(t_0)$ and $V_t(t_0, x_0) = H(t_0, x_0, \widehat{u}(t_0), \widehat{p}(t_0))$.

All necessary conditions of the Calculus of Variations are special cases of the Euler-Lagrange equation (and the transversality condition), after a suitable reformulation of the problem at hand.

3.3 Optimal Control: continuous time

We consider the Pontryagin problem. This is the problem that you get from the Lagrange problem when you: (1) add a constraint of the type $u(t) \in U \, \forall t \in [t_0, t_1]$ for some set $U \subseteq \mathbb{R}^r$ and (2) let the functions $x(\cdot)$ and $u(\cdot)$ run over larger function spaces: $x(\cdot)$ is allowed to be piecewise continuously differentiable (so it can have one or more kinks) and $u(\cdot)$ is allowed to be piecewise continuous (so it can have one or more jumps).

Theorem (Pontryagin's Maximum Principle). *Consider the Pontryagin problem given above. Conclusion: for each solution $(\widehat{x}(\cdot), \widehat{u}(\cdot))$ there is a number $\widehat{\lambda}_0 \geq 0$ and a C^1-function $\widehat{p}(\cdot) : [t_0, t_1] \to \mathbb{R}^n$, not both zero, such that (again in shorthand notation) $\dot{\widehat{x}} = \widehat{H}_p$, $\dot{\widehat{p}} = -\widehat{H}_x, \forall t \in [t_0, t_1]\, \widehat{H}(t) = \max_{u \in U} H(t, \widehat{x}(t), u, \widehat{p}(t), \widehat{\lambda}_0), \widehat{p}(t_1) = -\widehat{\kappa}'(\widehat{x}(t_1))$.*

The third condition gives PMP its name: it is called the maximum principle condition: it states that for each $t \in [t_0, t_1]$ the function $U \to \mathbb{R}$ given by $u \mapsto H(t, \widehat{x}(t), u, \widehat{p}(t), \widehat{\lambda}_0)$ assumes its maximal value at $u = \widehat{u}(t)$.

Here are two special cases where the maximum principle condition can be simplified:

- **Bang-bang.** If $U = [a, b] \subset \mathbb{R}$ and the Hamiltonian is linear in u: $H = A(t, x)u + B(t, x)$, then $\widehat{u}(t) = \text{bang}[a, b; A(t, x)]$. This is shorthand for $\widehat{u}(t) = a$ if $A(t, \widehat{x}(t)) < 0$ and $\widehat{u}(t) = b$ if $A(t, \widehat{x}(t)) > 0$.
- **Saturation.** If $U = [a, b] \subset \mathbb{R}$ and for each $t \in [t_0, t_1]$ the Hamiltonian $H(t, \widehat{x}(t), u, \widehat{p}(t), \widehat{\lambda}_0)$ is a quadratic function of u with graph a mountain parabola, say $A(t)u^2 + B(t)u + C(t)$ with $A(t) < 0 \forall t \in [t_0, t_1]$, then $\widehat{u}(t) = \text{sat}[-B(t)/2A(t); a, b]$. This is shorthand for: $\widehat{u}(t)$ is the point where the parabola has its maximum $-B(t)/2A(t)$, provided this is allowable (that is, lies in the set $U = [a, b]$); otherwise $\widehat{u}(t)$ is the boundary point of $[a, b]$ that is closest to $-B(t)/2A(t)$.

Numerical illustrations. The optimal policy in Example 12.2.4 is a bang-bang policy, the optimal policy in Example 12.2.6 is a saturation policy.

Convincing illustration. Newton's aerodynamic problem (see section 12.3.3 in [B-T]).

3.4 *Calculus of Variations: discrete time

We discretize the case that in the Pontryagin principle the state equation is $\dot{x} = u$: this gives $x_{k+1} = x_k + u_k$. That is, we take $\varphi(k, x, u) = x - u$. Moreover, we take $n = 1$ and $U = \mathbb{R}$. Finally we fix the end state x_N.

Theorem (Euler equations). *Given the problem*

$$\sum_{k=0}^{N-1} f(k, x_k, u_k) \to \min, x_k \in \mathbb{R}, 0 \leq k \leq N, u_k \in \mathbb{R}, 0 \leq k \leq N-1,$$

$$x_{k+1} = x_k + u_k,\ 0 \leq k \leq N-1, x_0 = a, x_N = b.$$

A necessary condition for optimality of a sequence of decisions is that this satisfies the Euler equations and transversality condition:

$$f_k(k, x_k, u_k) = f_u(k, x_k, u_k) - f_u(k-1, x_{k-1}, u_{k-1}), 0 \leq k \leq N-1,$$

$$f_u(N-1, x_{N-1}, u_{N-1}) = 0.$$

Comparing these conditions to the Euler differential equations and transversality condition given in lecture 3, we see that these conditions are the discretized versions, as it should be.

Sketch proof. Eliminate the decision variables: $u_k = x_{k+1} - x_k$. This gives the objective function $\sum_{k=0}^{N-1} f(k, x_k, x_{k+1} - x_k)$. Apply Fermat: put the partial derivatives equal to zero. This gives the Euler equations.

Remark. In the same way one can get a discrete version of the Euler-Lagrange equations.

3.5 *Optimal Control: discrete time

We take again the discrete version of the Pontryagin problem, but now we write the state equation as $x_{k+1} - x_k = g(k, x_k, u_k)$ instead of as $x_{k+1} = \varphi(k, x_k, u_k)$; that is, we take $g(k, x, u) = \varphi(k, x, u) - x$. Consider the Hamiltonian functions $H_k(x_k, u_k, p_k, \lambda_0) = -\lambda_0 f(k, x_k, u_k) + p_k g(k, x_k, u_k)$, for all $k = 0, \ldots, N-1$

Theorem (Pontryagin's maximum principle). *A necessary condition for optimality is given by the following equations, for a suitable nonzero collection of multipliers $\widehat{\lambda}_0, \widehat{p}_0, \ldots, \widehat{p}_{N-1}$.*

$$\widehat{x}_{k+1} - \widehat{x}_k = \frac{\partial \widehat{H}_k}{\partial p_k},$$

$$\widehat{p}_{k-1} - \widehat{p}_k = \frac{\partial \widehat{H}_k}{\partial x_k},$$

$$\widehat{H}_k = \max_{u_k \in U} J_k(\widehat{x}_k, u_k, \widehat{p}_k).$$

Again, if we compare these conditions to Pontryagin's maximum principle given in lecture 3, we see that these conditions are the discretized versions.

3.6 Problems on the lecture notes:

1. Give the details of the following step in the proof that the Euler equation is a necessary condition: $\int_{y_0}^{t_1} (\widehat{L}_x h + \widehat{L}_{\dot{x}} \dot{h}) dt = 0$ for a C^1 function $h : [t_0, t_1] \to \mathbb{R}$ for which $h(t_0) = h(t_1) = 0$ implies, by means of partial integration, $\int_{t_0}^{t_1} \langle \widehat{f}_{\dot{x}}(t) + \int_t^{t_1} \widehat{f}_x(\tau) d\tau, \dot{h}(t) \rangle dt = 0$.

2. Derive from Pontryagin's Maximum Principe the Euler equation and the transversality condition.

3. Derive from Pontryagin's Maximum Principle the bang-bang simplification. [Hint: use your knowledge of straight lines.]

4. Derive from Pontryagin's Maximum Principle the saturation simplification. [Hint: use your knowledge of parabolas.]

5. Consider the problem of finding the shortest connection between a point (t_0, x_0) with $t_0 \leq 0$ and a line (for which we take the vertical axis $t = 0$) by means of the graph of a C^1-function $x(\cdot)$. Formalize this problem as a Lagrange problem. Solve it by necessary conditions, for example by the Euler equation and the transversality condition.

3.7 Your tasks between lecture 3 and lecture 4.

You are expected to do as follows.

- In groups of two you have to solve the ten problems on the lecture notes, given above, as well as the following ten optimization problems (everything in your own words) by means of the four step method; you have to hand in your work on Tuesday 1-10-2013. Present complete solutions (not just some calculations). The problems 1-5 are from [B-T]:

  1. 3.6.10
  2. 4.2.1
  3. $\int_0^2 (6tx + \dot{x}^2)dt \to \min, x(0) = 0, x(2) = 10$ [hint: look at example 12.2.1]
  4. $\int_0^2 (6tx + \dot{x}^2)dt \to \min, x(0) = 0, x(2) = \text{free}$ [hint: look at example 12.2.2]
  5. $\int_0^1 u^2 dt \to \min, (x(\cdot), u(\cdot)) \in PC^1[0,1] \times PC[0,1], \dot{x} = 1 + u, x(0) = 0, x(1) = 2, u \geq 0.$

- Study these lecture notes carefully (if something is not clear, look it up in [B-T] or [K-S]).

- Study carefully in [B-T] the royal road as described on p.477 (ch 12).

- Study carefully in [B-T] the analysis of growth problem (see section 12.3.1).

- Study how the methods of optimization are applied to mathematical problems in chapter 9 (without worrying about the technical details).

- Try to look at the material of the chapters 9, 10,12 in [B-T] and try to look at the material in Kamien and Schwartz; conclude that the material from this nutshell covers a large part of that in [K-S].

- Try to solve exercises from [B-T] independently; in any case write the solutions in your own words.

4 Lecture on dynamic optimization II: Dynamic Programming (Bellman); introduction to infinite horizon stationary problems

4.1 Dynamic Programming: continuous time

We consider the Pontryagin problem ($P_{(t_0.x_0)}$) for each pair (t_0, x_0) with $t_0 \leq t_1$ consisting of a start time t_0 and an initial state x_0. That is, we consider a family of problems. We recall:

$$J(x(\cdot), u(\cdot)) = \int_{t_0}^{t_1} f(t, x(t), u(t))dt + \kappa(x(t_1)) \to \min, \dot{x} = \varphi(t, x(t), u(t)), x(t_0) = x_0, (x(t_1) = x_1), u \in U \tag{$P_{(t_0,x_0)}$}$$

Relation open and closed loop solution. Assume that the problem has a solution $(x_{t_0,x_0}(\cdot), u_{t_0,x_0}(\cdot))$—or, to be more precise, a family of solutions—for each pair (t_0, x_0). It is reasonable to restrict attention to those solutions that are consistent in the following sense: if for two pairs (t_0, x_0) and $(\bar{t}_0, \bar{x}_0)$ the functions $x_{t_0,x_0}(t)$ and $x_{\bar{t}_0,\bar{x}_0}(t)$ of t have the same value at some moment τ, then the functions $u_{t_0,x_0}(t)$ and $u_{\bar{t}_0,\bar{x}_0}(t)$ of t have the same value at all $t \geq \tau$ (and so $x_{t_0,x_0}(\cdot)$ and $x_{\bar{t}_0,\bar{x}_0}(\cdot)$ have the same value at all $t \geq \tau$). This is called a solution of the family of problems in open loop form. Now we can define $u(t_0, x_0) = u_{t_0,x_0}(t_0)$ for each pair (t_0, x_0) with $t_0 \leq t_1$. It is the best decision that can be taken in situation (t_0, x_0). Moreover, we define $V(t_0, x_0)$ to be the optimal value of the given problem for pair (t_0, x_0). As t_0 and x_0 are arbitrary (with the restriction $t_0 \leq t_1$ of course), we can omit the subscripts 0: each t can be viewed as starting time and each x can be viewed as initial state. Thus we write $u(t, x)$ and $V(t, x)$. Then the pair of functions $(u(t, x), V(t, x))$ of t and x is called a solution of the family in closed loop form; $(t, x) \mapsto u(t, x)$ is called an optimal policy in closed loop form, also called optimal policy in feedback form.

Switching between open and closed loop form. 1) Given a solution in closed loop form $(u(t, x), V(t, x))$. Then, for each (t_0, x_0), the function $x_{t_0,x_0}(t)$ of t is defined to be the solution of the Cauchy problem $\dot{x}(t) = \varphi(t, x(t), u(t, x(t))), x(t_0) = x_0$ and $u_{t_0,x_0}(\cdot)$ is defined to be the function $u(t, x_{t_0,x_0}(t))$ of t. 2) Given a solution in open loop form $(x_{t_0,x_0}(\cdot), u_{t_0,x_0}(\cdot))$. Then $u(t_0, x_0)$ is defined to be $u_{t_0,x_0}(t_0)$ and $V(t_0, x_0)$ is defined to be $J(x_{t_0,x_0}(\cdot), u_{t_0,x_0}(\cdot))$. The we drop again the subscripts 0 and write $u(t, x)$ and $V(t, x)$.

Theorem (Bellman equation). *Assume that $V(t, x)$, the optimal value function of the problem, is C^1. Then*

$$-V_t(t, x) = \inf_{u \in U}[f(t, x, u) + V_x(t, x)\varphi(t, x, u)]$$

for all (t, x), and $V(t_1, x) = \kappa(x) \forall x \in \mathbb{R}$. The value of u for which the infimum is assumed is the optimal decision $u(t, x)$ in situation (t, x). These properties of the pair of functions $(u(t,x), V(t, x))$

are also sufficient for the pair to be a solution of the family in closed loop form.

Sketch of the proof. We assume that $f \equiv 0$ (This can be achieved by reformulation of the problem. To this end one increases the dimension of the state by one, by introducing the state variable $x_{n+1}(t)$ satisfying $\dot{x}_{n+1} = f(t, x(t), u(t))$ and $x_{n+1}(t_0) = 0$ and we change the objective function ito $x_{n+1}(t_1) + \kappa(x(t_1)$.) Then one has for each immediate decision u in situation (t, x) that $V(t, x) \leq V(t + dt, x + dx)$ if $dx = \varphi(t, x, u)dt$, and equality holds precisely if u is the optimal decision in situation (t, x). Now use that $V(t+dt, x+dx) = V(t, x) + V_t(t, x)dt + V_x(t, x)dx$ and that $dx = \varphi(t, x, u)dt$ and this gives the Bellman equation.

Remark. The proof shows that, in the case $f \equiv 0$, for each admissible $(x(\cdot), u(\cdot))$ the function $t \mapsto V(t, x(t))$ is monotonic non-decreasing; it is constant precisely if $(x(\cdot), u(\cdot))$ is optimal.

Illustration. Boat story (section 12.2.4).

How to make use of the Bellman equation. There are three possibilities: 1) sometimes you do no more than writing down and simplifying the Bellman equation: this gives some insight, 2) sometimes you can make an educated guess for the form of the optimal value function $V(t, x)$; only some parameters in this form have to be determined. Substitution of this attempt in the Bellman equation then leads to a system of equations in these parameters, from which these can be calculated; thus $V(t, x)$ is calculated and then $u(t, x)$ can be calculated, as the value of u for which the expression $f(t, x) + V_x(t, x)\varphi(t, x, u)$ takes a minimum, 3) you can use the Bellman equation as a verification method. First you 'solve' the problem (with the parameters t_0, x_0 in it) by means of necessary conditions, by Optimal Control (Pontryagin's Maximum Principle) or one of its weaker but easier variants from the Calculus of Variations (Euler, transversality, Euler-Lagrange). This leads to a candidate solution for each problem of the family of Pontryagin problems in open loop form. That this is a candidate for the *family* means in practice that the formulas for $x(\cdot)$ and $u(\cdot)$ involve the parameters t_0, x_0. Now you switch to closed loop form and then you verify that these candidates in closed loop form satisfy the Bellman equation. This proves that the candidates in open loop form are solutions of the given family of Pontryagin problems. This is a very powerful method. However, the verification involves a lot of formula manipulation!

4.2 Dynamic Programming: discrete time

Here is the result of discretizing the Pontryagin problem:

$$J((x_k), (u_k)) = (\sum_{k=0}^{N-1} f(k, x_k, u_k)) + \kappa(x_N) \to \min, x_k \in \mathbb{R}, 0 \leq k \leq N, u_k \in U, 0 \leq k \leq N-1,$$

$$x_{k+1} = \varphi(k, x_k, u_k), 0 \leq k \leq N-1, x_0 = a.$$

Here $U \subseteq \mathbb{R}^m$, $f : \{0, 1, \ldots, N-1\} \times \mathbb{R}^n \times U \to \mathbb{R}$, $\kappa : \mathbb{R}^n \to \mathbb{R}, \varphi : \{0, 1, \ldots, N-1\} \times \mathbb{R}^n \times U \to \mathbb{R}^n$. Here starting time is 0 and initial state is a; we denote the problem as $(P_{0,a})$. We embed this problem again into a family of problems by considering it for all initial pairs ('situations') $(\bar{k}, \bar{x})$, that is, starting time $\bar{k}$ and initial state $\bar{x}$. We denote the problem with initial situation $(\bar{k}, \bar{x})$ as $(P_{\bar{k},\bar{x}})$. We define a policy in closed form (or in feedback form) to be a function $\pi : \{0, \ldots, N-1\} \times \mathbb{R}^n \to \mathbb{R}^m$. One can carry out this (not necessarily optimal) policy and this leads to a value function $V : \{0, \ldots, N\} \times \mathbb{R}^n \to \mathbb{R}$: substitution of the policy in the state equation gives $x_{k+1} = \varphi(k, x_k, \pi(k, x_k))$; this can be solved recursively, starting with situation $(\bar{k}, \bar{x})$; this gives the time path $x_k, \bar{k} \leq k \leq N$; substitution in the policy function gives then the time path of the decisions $u_k = \pi(k, x_k), \bar{k} \leq k \leq N-1$; substituting these two time paths into the objective function gives the value $V(\bar{k}, \bar{x})$ that you get if you carry out the given policy in closed loop form starting from situation $(\bar{k}, \bar{x})$. The policy is called optimal if for each pair $(\bar{k}, \bar{x})$ the value of this policy equals the optimal value $V^*(\bar{k}, \bar{x})$ of the problem with initial situation $(\bar{k}, \bar{x})$.

Theorem. *One has the following criterion for the optimal value function V^*: it has to satisfy the functional equation $V(k, x) = \inf_{u \in U}[f(k, x, u) + V(k+1, \varphi(k, x, u))]$, $0 \leq k \leq N-1, x \in \mathbb{R}^n$, and the boundary condition $V(N, x) = \kappa(x), x \in \mathbb{R}^n$. Moreover, a policy π is optimal if and only if the infimum is assumed in $\pi(k, x)$ for all $k \in \{0, \ldots, N-1\}$, $x \in \mathbb{R}$.*

Remark. This functional equation with boundary condition can be solved recursively, with time running backwards: $V(k, x)$ is determined for $k = N$ by the boundary condition, and the functional equation determines $V(k, x)\, \forall x$ if $V(k+1, x)\, \forall x$ is known. In particular, there exists a solution for the functional equation with boundary condition and this solution is unique.

Proof. The optimal value $V^*(k, x)$ is by definition the infimum over all sequences of controls $u_k, \ldots, u_{N-1}$ of the objective function for the problem with initial situation (k, x). One can carry out this minimization process in two steps: first one takes the infimum over all sequences of controls $u_k, \ldots, u_{N-1}$ with u_k fixed in an arbitrary way, say $u_k = u$ for some $u \in U$: this gives $f(k, x, u) + V^*(k+1, \varphi(k, x, u))$. Then one takes the infimum of this outcome over all $u \in U$. This give $\inf_{u \in U}(f(k, x, u) + V^*(k+1, \varphi(k, x, u))$; on the other hand this gives $V^*(k, x)$. Thus we get the Bellmann equation. This argument shows also the sufficiency and the statement about the optimal policy in the theorem.

Fixed point interpretation. We can define an operator T on functions $V : \{0, \ldots, N\} \times \mathbb{R}^n \to \mathbb{R}$ by putting $(TV)(k, x) = \inf_{u \in U}[f(k, x, u) + V(k+1, \varphi(k, x, u))]$ for all $0 \leq k \leq N-1, x \in \mathbb{R}^n$ and $(TV)(N, x) = \kappa(x), x \in \mathbb{R}^n$. Then a solution of the Bellmann equation with boundary condition is the same thing as a fixed point of the operator T, $V = TV$. This suggests the following algorithm: start with an arbitrary function $V(k, x)$. Then compute successively TV, T^2V, T^NV. Then T^NV is the optimal value function V^*. The interpretation of T^lV is: it is the optimal value if you can make decisions u_k only at moments $k = 0, 1, \ldots, l-1$, in this way letting the state evolve from $x_0 = a$ to

x_l, and incurring costs $f(k, x_k, u_k)$ at times $k = 0, \dots, l-1$; from moment l the cost is prescribed to be $V(l, x_l)$.

Remark. It is of interest to focus on the case that $f \equiv 0$ (to which we can always reduce by reformulation of the problem). Then the Bellman equation reduces to $V(k, x_k) \le V(k+1, x_{k+1})$ for all $u \in U, x_k \in \mathbb{R}^n$ and with $x_{k+1} = \varphi(k, x_k, u)$, and equality holds iff u is the optimal decision in situation (k, x_k). Then, if we carry out any policy, we get that the value function $V(k, x_k)$ is monotonic non-decreasing; it is constant iff the policy is optimal.

Illustration. Tower of Pisa (section 11.3.4).

4.3 Stationary infinite horizon problems

We consider the following 'social planner'-type problem:

$$I(x(\cdot)) = \sum_{t=0}^{\infty} \beta^t U(x(t), x(t+1)) \to \max;\ x(\cdot) : \mathbb{N}_0 \to X, x(t+1) \in G(x(t))\, \forall t \ge 0, x(0) = x_0. \quad (P)$$

Here $\beta \in (0,1)$, $X \subseteq \mathbb{R}^n$, $G(x)$ is a nonempty compact (that is, closed and bounded) set in X $\forall x \in X$, $U : X_G \to \mathbb{R}$ is continuous and bounded, where $X_G = \{(x,y) \in X \times X | y \in G(x)\}$ and $x_0 \in X$.

Social planner problem viewed as a discrete version of the Pontryagin problem. The social planner problem can be viewed as a special case of the discrete version of the Pontryagin problem. Assume that $f(k, x_k, u_k)$ only depends directly on time k by discounting: $f(k, x_k, u_k) = \beta^k g(x_k, u_k)$ for some $\beta \in (0,1)$; moreover assume that $\varphi(k, x_k, u_k)$ does not depend on k, so that we can write $x_{k+1} = \varphi(x_k, u_k)$. Assume that for each pair (k, x_k) the value $f(k, x_k, u_k)$ does not depend on u_k but only on x_{k+1}, the next state that is reached by decision u_k, that is $x_{k+1} = \varphi(x_k, u_k)$. (This is for example the case if one can rewrite the state equation $x_{k+1} = \varphi(x_k, u_k)$ in such a way that u_k is expressed in k and x_k.) Then $f(k, x_k, u_k)$ can be written as $\beta^k g(x_k, x_{k+1})$. Write G_k for the set of states that can be reached from state x_k by a suitable decision in U, that is, $G_k = \{\varphi(k, x_k, u) | u \in U\}$. Then one can simplify the determination of a next state ('first make a decision, then this leads to a new state') to: given state x_k at time k, the decision is to choose a next state x_{k+1} in G. Furthermore, we work with infinite horizon instead of with final time N. Finally we write the problem as a maximization problem in the usual way ('minus sign') and write $U = -g$. Then we get the social planner problem.

We embed this problem into the following family of problems $(P_{\tau,\xi})$, with $\tau \in \mathbb{N}_0, \xi \in X$:

$$I_\tau(x(\cdot)) = \sum_{t=\tau}^{\infty} \beta^t U(x(t), x(t+1)) \to \max; x(\cdot) : \{\tau, \tau+1, \ldots\} \to X, x(t+1) \in G(x(t)) \, \forall t \geq \tau, x(\tau) = \xi \qquad (P_{\tau,\xi}).$$

Let $\tilde{V}^*(\tau, \xi)$ be the optimal value of $(P_{\tau,\xi})$ and let $\tilde{M}(\tau, \xi)$ be the set of solutions of $(P_{\tau,\xi})$.

Stationarity. Problem $(P_{\tau,\xi})$ is the same as problem $(P_{0,\xi})$ up to discounting: by multiplying the objective function of $(P_{0,\xi})$ with β^τ, one gets problem $(P_{\tau,\xi})$. In particular, $\tilde{V}^*(\tau, \xi) = \beta^\tau \tilde{V}^*(0, \xi)$ and $M(\tau, \xi) = M(0, \xi)$. We write $V^*(\xi) = \tilde{V}^*(0, \xi)$ and $M(\xi) = \tilde{M}(0, \xi)$ for all $\xi \in X$. We restrict attention to the problems $P_\xi = P_{0,\xi}$. We have

$$I(x(\cdot)) \to \max, x(\cdot) \in \Pi(\xi) \qquad (P_\xi)$$

where $\Pi(\xi) = \{x(\cdot) : \mathbb{N}_0 \to X | x(0) = \xi, x(t+1) \in G(x(t)) \forall t \geq 0\}$ for each $\xi \in X$.

Policies in closed and open loop form. A policy in closed loop form is a mapping $\pi : X \to X$ such that $\pi(\xi) \in G(\xi) \forall \xi \in X$. A policy in open loop form—not necessarily an optimal one—is a collection $x^\xi(\cdot) \in \Pi(\xi)$ $\forall \xi \in X$ for which the following consistency property holds: $x^\xi(s) = x^\eta(t) \Rightarrow x^\xi(s+1) = x^\eta(t+1)$ $\forall \xi, \eta \in X, \forall s, t \geq 0$. Policies in closed and open loop form are equivalent. To switch from a policy in closed loop form π to a policy in open loop form $x^\xi(\cdot), \xi \in X$, one defines $x^\xi(\cdot)$ recursively for all $\xi \in X$ by: $x^\xi(0) = \xi, x^\xi(t+1) = \pi(x^\xi(t)) \, \forall t \geq 0$. To switch from a policy in open loop form $x^\xi(\cdot), \xi \in X$ to a policy in closed loop form, one defines π by $\pi(\xi) = x^\xi(1)$ for all $\xi \in X$.

Bellman equation with supplement and boundary condition. For a pair (V, π) consisting of a function $V : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and a policy in closed loop form π, we consider the following equations:

1. the Bellman equation: $V(\xi) = \sup_{\eta \in G(\xi)}(U(\xi, \eta) + \beta V(\eta))$ $\forall \xi \in X$ or, in fixed point form, $V = TV$ with $(TV)(\xi) = \sup_{\eta \in G(\xi)}(U(\xi, \eta) + \beta V(\eta))$ $\forall \xi \in X$ (functional equation in the value function V),
2. the supplement Bellman equation: the infimum above is assumed in $\pi(\xi)$ for all $\xi \in X$ (characterization optimal policy π if V is given).
3. the boundary condition $\lim_{t\to\infty} \beta^t V(t) = 0$.

Theorem (The Bellman equation as criterion for optimality). *(1) The optimal value function V^* is the unique solution to the Bellman equation with boundary condition. (2) An admissible element $\widehat{x}(\cdot)$ for problem (P) is optimal iff $V^*(\widehat{x}(t)) = U(\widehat{x}(t), \widehat{x}(t+1)) + \beta V^*(\widehat{x}(t+1))$ $\forall t \geq 0$.*

Remark. More generally, one has that for a policy in open loop form $x^{\xi}(\cdot), \xi \in X$, the time path $x^{\xi}(\cdot)$ is a solution of (P_{ξ}) for all $\xi \in X$ iff the corresponding policy in closed loop form π satisfies $V^*(\xi) = U(\xi, \pi(\xi)) + \beta V^*(\pi(\xi))$ for all $\xi \in X$.

Proof. (1) Choose $\xi \in X$. $V^*(\xi)$ is the supremum of the expression $\sum_{k=0}^{\infty} \beta^t U(x(t), x(t+1))$ over all $x(\cdot) \in \Pi(\xi)$; this can be calculated by taking first the supremum over all $x(\cdot) \in \Pi(\xi)$ with $x(1)$ fixed, say $x(1) = \eta \in G(\xi)$—this gives $U(\xi, \eta) + \tilde{V}^*(1, \eta)$ and we know that $\tilde{V}^*(1, \eta) = \beta V^*(\eta)$—and then taking the supremum of the outcome over all $\eta \in G(\xi)$. This shows that V^* satisfies the Bellman equation. As U is bounded, the boundary condition is also satisfied. Now let V be an arbitrary solution of the Bellman equation and the boundary condition. Then for each $\xi \in X$ and each $k \in \mathbb{N}$, we have $V(\xi) = (T^k V)(\xi) = \sup_{x(\cdot) \in \Pi(\xi)} [\sum_{t=0}^{k} \beta^t U(x(t), x(t+1))] + \beta^{k+1} V(x((k+1))$. By the boundary condition, it follows on taking $k \to \infty$ that $V(\xi) = \sup_{\xi(\cdot) \in \Pi(\xi)} \sum_{t=0}^{\infty} \beta^t U(x(t), u(t+1)) = V^*(\xi)$. Therefore, $V = V^*$.

(2) If $\widehat{x}(\cdot) \in M(x_0)$, then $\widehat{x}(\cdot)$ restricted to $\{s, s+1, \ldots\}$ is in $M(\widehat{x}(s))$ for each $s \in \mathbb{N}$ by the principle of optimality. Therefore, $V^*(\widehat{x}(t)) = \sum_{k=t}^{\infty} \beta^{k-t} U(\widehat{x}(k), \widehat{x}(k+1))$ and $V^*(\widehat{x}(t+1)) = \sum_{k=t+1}^{\infty} \beta^{k-t-1} U(\widehat{x}(k), \widehat{x}(k+1))$ and so $V^*(\widehat{x}(t)) = U(\widehat{x}(t), \widehat{x}(t+1)) + \beta V^*(\widehat{x}(t+1)))$. Conversely if $V^*(\widehat{x}(t)) = U(\widehat{x}(t), \widehat{x}(t+1)) + \beta V^*(\widehat{x}(t+1)))\ \forall t \geq 0$, then starting with this equation for $t = 0$ and then substituting successively in it the equations for $t = 1, \ldots, k$ and then taking the limit $k \to \infty$ gives $V^*(x_0) = \sum_{t=0}^{\infty} U(\widehat{x}(t), \widehat{x}(t+1))$; that is, $\widehat{x}(\cdot) \in M(x_0)$.

Convincing application. Social planner problem (see Acemoglu).

Problems on the lecture notes.

1. In the previous assignment you solved the necessary conditions for the problem of finding the shortest connection between a point (t_0, x_0) with $t_0 \leq 0$ and a line (for which we take the vertical axis $t = 0$) by means of the graph of a C^1-function $x(\cdot)$. Now switch from the candidate solution in open loop form that you have found to closed loop form.

2. Check that the closed form candidate that you have found in the previous problem satisfies the Bellman equation. What is the use of this?

3. Explain why in the family of Pontryagin problems in the special case $f \equiv 0$, the function $t \mapsto V(t, x(t))$ is monotonically non-decreasing for each admissible pair $(x(\cdot), u(\cdot))$ for one problem of the family. Explain moreover why this function $t \mapsto V(t, x(t))$ is constant precisely if the pair $(x(\cdot), u(\cdot))$ is optimal [Hint: use the language of the boat story.]

4. Give an interpretation for $T^2 V$ in the finite horizon case and prove this interpretation.

5. With the Bellman approach you always have to consider the problem for all future moments as initial moment. However, for the stationary infinite horizon problem we only consider initial

time $t = 0$. How is this possible?

Your tasks between lecture 4 and lecture 5. You are expected to do as follows.

- In groups of two you make the ten problems on the lecture notes given above, as well as the following five problems. You have to hand in your work on time. Present complete solutions.

  1. 3.6.5 [hint: do not try to calculate the multipliers, but try to find out which information the Lagrange equations give on the optimum]
  2. Solve the tower of Pisa problem by means of the Bellman equation (see exercise 4.6.12).
  3. $\int_0^1 (\dot{x}^2 + x^2)dt \to \min, x \in PC^1[0,1], x(0) = 2$ (hint: write u for $\dot{x}$ and use the Euler equation and the transversality equation).
  4. 11.4.7
  5. *Give the derivation of the remark after the theorem on the Bellman equation (infinite horizon) from this theorem.

- Study these lecture notes carefully
- Study the royal road that is given on p 443 (ch 11)
- Try to look at chapter 11 and at sections 12.4.1 and 12.4.2.
- Here are some additional questions on the lecture notes (not to be handed in; maybe you can look at these problems):

  1. Write out explicitly what problem $P_{\bar{k},\bar{x}}$ is.
  2. Can it happen that the Bellman equation plus boundary condition for finite horizon have more than one solution?
  3. Give in your own words a description of the Bellman equation for finite horizon and of the reason that it holds.
  4. How can you reduce to the case $f \equiv 0$ in the finite horizon case?
  5. Give a direct and simple proof for the Bellman equation in the case $f \equiv 0$.
  6. Explain in detail the intuition behind the boundary condition of the Bellman equation for infinite horizon.

- Here are some additional problems (not to be handed in; maybe you can look at these problems, especially at the last one)

1. $\int_0^1 u^2 dt \to \min, (x(\cdot), u(\cdot)) \in PC^1[0,1] \times PC[0,1], \dot{x} = u, x(1) = 3, u \geq 0$; here you have to use the Bellman equation. (hint: work out the Bellman equation in this case and then write down the necessary conditions for the optimization that is given by the right hand side of the Bellman equation ('$\inf_{u \in U}$'). Use these to simplify the Bellman equation. This is sufficient in order to get full points. You can go on, for the honor, to try and solve the Bellman equation.)
2. 4.6.13
3. 11.3.1
4. 11.4.1
5. Explain briefly the following slogans that are sometimes used in the context of dynamic programming:
   - you make the problem easier by making it more difficult
   - each day is the first day of the rest of your life
   - the domino effect: one problem is easy to solve; the solution helps to solve a neighboring problem; going on in this way, you come to the solution of the given problem
   - embed a problem into a family of problems and use the relation between the solutions of neighboring problems
   - each tail of an optimal plan is optimal
   - a multi-decision problem is turned into several single decision problems
   - you should use recursion/backward induction
   - solve by induction with respect to 'time to go'
   - the minimal cost that is achievable from some starting point equals the minimum of the sum of the cost of the immediate decision plus the minimal cost that is achievable from the point that you reach by this decision; here you minimize over all possible immediate decisions
   - if you are by mistake not on the optimal route anymore, you should not look back in regret over the mistake; instead you should only look forward and make the best of the remaining time by seeing your current situation as the starting point of a new problem
   - you should make an optimal plan for all possible future contingencies rather than find one optimal solution to one dynamic optimization problem
   - find a feedback/closed loop solution rather than an open loop solution

5 Lecture on dynamic optimization III: : infinite horizon stationary problems

We are going to use the fixed point interpretation of the Bellman equation with boundary condition to obtain information about its solution, for example its continuity properties.

5.1 Contraction mapping theorem

The contraction mapping theorem is the fundamental existence and uniqueness result for fixed points.

Metric space. A metric space is a pair (S, d) consisting of a set S and a function $d : S \times S \to [0, \infty)$ for which (1) $d(s, u) \leq d(s, t)+d(t, u)\,\forall s, t, u \in S$, (2) $d(s, t) = d(t, s)\,\forall s, t \in S$, (3) $d(s, t) = 0 \Leftrightarrow s = t$. Then $d(s, t)$ is called the distance between s and t.

Limits in a metric space. An infinite sequence $\{s_k\}$ in a metric space (S, d) has limit $s \in S$ if $\lim_{k\to\infty} d(s_k, s) = 0$. The limit is unique if it exists.

Cauchy property sequence. An infinite sequence $\{s_k\}$ for which $d(s_k, s_l)$ tends to zero if $k, l \to \infty$—to be more precise, $\lim_{n\to\infty} \sup_{p>0} d(s_{n+p}, s_n) = 0$—is said to have the Cauchy property. Then it is called a Cauchy sequence.

Internal necessary condition for existence limit. If an infinite sequence has a limit, then it has the Cauchy property. But the converse need not be true: take as metric space S the real line with the number zero removed: $\mathbb{R} \setminus \{0\}$ ('the line with a hole in it'); take the usual distance $d(x, y) = |x - y|$. Then the sequence $s_n = 1/n$ is a Cauchy sequence in S that does not have a limit (in $\mathbb{R}$ it has a limit, 0, but this point is not in S, as we removed it).

Completeness (= 'The Cauchy property is an internal criterion for having a limit'). A metric space is called complete if the Cauchy property is a criterion for sequences to have a limit. So in complete spaces one can show that sequences have a limit without the need to produce this limit: you just have to check that it has the Cauchy property, which is an internal property: you do not have to go outside the sequence to verify it.

Example 1. The metric space $S = \mathbb{R}$ with distance $d(x, y) = |x - y|$, the absolute value of the difference, for all $x, y \in \mathbb{R}$, is complete.

Example 2. The metric space $S = \mathbb{R}^n$ with distance $d(x, y) = |x - y| = \sqrt{\sum_{k=1}^{n}(x_k - y_k)^2}$, the euclidean distance, for all $x, y \in \mathbb{R}^n$, is complete.

Example 3. Let X be a set. The metric space $S = B(X)$, consisting of the bounded functions $X \to \mathbb{R}$, with distance $d(f, g) = \|f - g\|_\infty = \sup_{x\in X} |f(x) - g(x)|$ for all bounded functions $f, g : X \to \mathbb{R}$, is a complete metric space.

These are the basic examples (in fact the examples 1 and 2 can be viewed as special cases of example 3: if we take in example 3 a set X of n elements, then $S = B(X)$ is essentially $\mathbb{R}^n$. Moreover, example 1 is the special case $n = 1$ of example 1). The following result is useful to make from old examples of complete metric spaces new examples.

Closed subsets. A subset T of a metric space S is closed if for each infinite sequence in T that has a limit ('is convergent') in S, this limit lies in T.

Proposition. *A closed subset of a complete metric space is complete, when viewed as a metric space by itself.*

Here is an example how this result can be used to make from old examples of complete metric spaces new ones.

Example 4. Let X be a subset of $\mathbb{R}^n$. The metric space $S = C(X)$, consisting of the bounded continuous functions $X \to \mathbb{R}$, with distance $d(f,g) = \sup_{x\in X} |f(x) - g(x)|$ is a complete metric space. This can be proved by showing that $C(X)$ is closed in $B(X)$.

Contraction mapping theorem. *Let (S,d) be a complete metric space and let $T : S \to S$ be a mapping for which there exists $\beta \in (0,1)$ with $d(Tz_1, Tz_2) \leq \beta d(z_1, z_2)$ for all $z_1, z_2 \in S$, then for each $\bar{s} \in S$, the sequence that is defined recursively by $s_{k+1} = Ts_k \, \forall k \geq 0, s_0 = \bar{s}$ has a limit s. This limit is independent of the starting point $\bar{s}$ and it is the unique fixed point of T, $Ts = s$.*

Proof. For any k, l, say with $k \leq l$, one has $d(s_k, s_l) \leq d(s_k, s_{k+1}) + \cdots + d(s_{l-1}, s_l) \leq (\beta^k + \cdots + \beta^{l-1}) d(s_1, s_0) \leq \frac{\beta^k}{1-\beta} d(s_0, s_1)$. This shows that the Cauchy property holds, and so by completeness, the sequence has a limit s. We have $d(s, Ts) \leq d(s, T^n s) + d(Ts, T^n s)| \leq d(s, T^n s) + \beta d(s, T^{n-1} s) \to 0$ for $n \to \infty$, so $s = Ts$. If u, v are fixed points, then $d(u,v) = d(Su, Sv) \leq \beta d(u,v)$, so $u = v$.

Theorem (Blackwell's sufficient condition for a contraction). *Given a set X, a subset $F \subseteq B(X)$ and an operator $T : F \to F$ for which the following two conditions hold:*

1. *$Tf \leq Tg$ for all $f, g \in F$ with $f \leq g$.*
2. *there exists $\beta \in (0,1)$ such that $T(f + c) \leq Tf + \beta c$ for all $f \in F, c \in [0, \infty)$.*

Then $d(Tf, Tg) \leq \beta d(f,g)$ for all $f, g \in F$.

Proof. Choose $f, g \in F$. One has $f \leq g + \|f - g\|_\infty$ and so, by conditions 1. and 2. $Tf \leq T(g + \|f - g\|_\infty) \leq Tg + \beta \|f - g\|_\infty$, that is, $Tf - Tg \leq \beta \|f - g\|_\infty$. By symmetry, one has $Tg - Tf \leq \beta \|g - f\|_\infty$ as well. This gives $\|Tf - Tg\|_\infty \leq \beta \|f - g\|_\infty$ as required.

Illustration. We can use this to prove (again) that the Bellman equation for the infinite horizon problem has a unique bounded solution: we know this already: V^* is a bounded function and it is the unique solution of the Bellman equation and its boundary condition, which is trivially satisfied for all bounded functions). We verify Blackwell's sufficient conditions for the operator $T : B(X) \to B(X)$ given by $(TV)(\xi) = \sup_{\eta \in G(\xi)}(U(\xi, \eta) + \beta V(\eta))$: (1) if $V_1 \leq V_2$, then $TV_1 \leq TV_2$, clearly, (2) $(T(V + c))(\xi) = \sup_{\eta \in G(\xi)}(U(\xi, \eta) + \beta(V(\eta) + c)) = (TV)(\xi) + \beta c$. Therefore, T is a contraction on $B(X)$ and so, the contraction mapping theorem gives that it has a unique fixed point.

The real application of the contraction mapping theorem. So far nothing new. Now come two interesting tricks, which will allow us to get truly new information by using the contraction mapping theorem. To wit, these allow us to establish that V^* possesses certain interesting properties.

1. If we have checked that the set P of functions V having a certain property is closed in $B(X)$ and that this set is left invariant under T (that is, $V \in P \Rightarrow TV \in P$) then it follows that V^* has the property.
2. If, moreover, for each V with the property under consideration, TV has a certain stronger property, and the set of functions having this stronger property is not closed, then it follows nevertheless that V^* has this stronger property.

Kindly study the following subtle proof carefully.

Proof.

1. We recall that $B(X)$ is complete. It is assumed that P is closed in $B(X)$, so we get by the proposition above that P is complete. Moreover, it is assumed that T restricts to a function $P \to P$; this is of course again a contraction. Now we apply the contraction mapping theorem. It gives that there exists a unique $V \in P$ with $TV = V$. We know already that V^* is the unique fixed point of T in $B(X)$. As $P \subset B(X)$, we get that $V^* \in P$!
2. By the argument above, V^* has the property, and so, by assumption TV^* has the stronger property. As $TV^* = V^*$, we get that V^* has the stronger property!!

5.2 The Maximum Theorem

The maximum theorem is the fundamental theorem on maximization problems that depending on a parameter, giving the continuity of the dependence of its value and its solution set on the parameter.

We begin by formulating the concept 'family of maximization problems' (or 'maximization problem depending on a parameter') in the general context of metric spaces.

Product of metric spaces. The product of two metric spaces (X, d_X) and (Y, d_Y) is again a metric space $(X \times Y, d_{X\times Y})$, with $d_{X\times Y}((x,y),(x',y')) = d(x,y) + d(x',y')\,\forall x,x' \in X, y,y' \in Y$.

Continuous function. For two given metric spaces (X,d), a mapping $f : X \to Y$ is called continuous if $\lim_{k\to\infty} f(x_k) = f(\lim_{k\to\infty} x_k)$ for each convergent sequence $\{x_k\}$ in X. This can be applied to functions $X \to \mathbb{R}$, viewing $\mathbb{R}$ as a metric space with distance function $d(x,x') = |x - x'|$.

Compactness. A metric space is called compact if each sequence has a convergent subsequence.

Criterion. Each compact metric space is complete and bounded. For subsets of $\mathbb{R}^n$ the converse holds, and this gives the following criterion: a subset of $\mathbb{R}^n$ is compact iff it is bounded and closed. We will not consider the property compactness for other metric spaces (for the record, we mention that there is a criterion for compactness for general metric spaces, which involves replacing boundedness in the criterion above by a variant of it: totally boundedness).

Theorem of Weierstrass. *A continuous function on a nonempty compact metric space assumes its minimum and maximum.*

Continuous functions. Let X, Y be metric spaces and $f : X \to Y$ a function. Then f is called continuous if for each convergent infinite sequence $x_1, x_2, \ldots$ in X, with limit $a \in X$, one has $\lim_{n\to\infty} f(x_n) = f(a)$.

Family of problems. We consider a family of continuous maximization problems $(P_y)_y$, where the index y runs over a metric space Y and the admissible set $\Gamma(y)$ of (P_y) is a nonempty compact subset of a given metric space X. To be more precise: let X and Y be metric spaces, $\Gamma : Y \to c(X)$ be a correspondence (that is a multivalued mapping, so it associates to each element of Y a subset of X and not an element of x!) where $c(X)$ is the collection of nonempty compact subsets of X, and $f : X \times Y \to \mathbb{R}$ be a continuous function; consider the family $(P_y), y \in Y$ with

$$f(x,y) \to \max, x \in \Gamma(y) \qquad (P_y)$$

Optimal value function and solution correspondence. For a family of continuous maximization problems $(P_y), y \in Y$, the optimal value function is denoted by $V^* : Y \to \mathbb{R}$ and the solution correspondence is denoted by $G : Y \to c(X)$, so $G(y)$ is the set of all global solutions of (P_y).

Proof that $G(y)$ is compact for all $y \in Y$. By definition, $G(y) = \{x \in \Gamma(y) | f(x,y) = V^*(y)\}$. Choose an infinite sequence $x_1, x_2, \ldots$ in $G(y)$. As $G(y) \subset \Gamma(y)$ and $\Gamma(y)$ is compact, there exists a subsequence $x_{i_1}, x_{i_2}, \ldots$ that is convergent in $\Gamma(y)$; call the limit u, so $u \in \Gamma(y)$. AS $x_{i_k} \in G(y)$, we have $f(x_{i_k}, y) = V^*(y)$ for all k. Take the limit $k \to \infty$, using that f is continuous, and then you get $f(u,y) = V^*(y)$. That is $u \in G(y)$. This finishes the proof that $G(y)$ is compact.

The question. Does continuity of Γ imply continuity of V^* and G? This is a natural question: does a small change in the parameter in the problem to a small change in its optimal value and its set of global solutions. To pose this question in a precise way requires a concept of continuity for the correspondences Γ and G. Such a concept exists; we give it below. It makes use of the fact that the set of compact subsets of a metric space form a metric space. This is all we need, as we have already a continuity concept for functions from one metric space to another. The definition of the distance function on the compact subsets of a metric space, called the Hausdorff metric, is unfortunately not as simple as we would like. It looks quite intricate and seems to pop up from nowhere. However, it is very desirable to have some concept of continuity here, and this is the simplest *conceptual* definition. Later we will make the point that it can be replaced by a simpler description; this is what is usually done in practice; then you do not need the Hausdorff metric.

Hausdorff metric. The set $c(X)$ of compact subsets of a metric space X can be made into a metric space. The points of this metric space are the elements of $c(X)$, that is, compact subsets of X; the metric on $c(X)$ is defined as follows. If $A, B \in c(X)$, then we define $d_H(A,B) = \max(\{\omega(A,B), \omega(B,A)\}$, where $\omega(A,B) = \sup\{d_X(z,B) : z \in A\}$ with $d_X(z,B) = \inf\{d_X(z,b) | b \in B\}$. Thus the concept continuity is defined for correspondences $Y \to c(X)$ that associate to each element of Y a compact subset of X (as we have a concept continuity for functions from one metric space to another one).

A disappointing answer. After all this trouble, it might come as an unpleasant surprise that the answer to the question above is negative. However, if we modify the question slightly, the answer is fortunately positive: it is the maximum theorem given below. There is a price to be paid here: the concept Hausdorff continuity has to be split up into two properties, upper hemicontinuous and lower hemicontinuous. We will display the definitions, but giving the intuition for these concepts falls outside the scope of this course (see Acemoglu for the intuition).

Upper hemicontinuity for a compact-valued correspondence. A mapping $\Gamma : Y \to c(X)$ is called a upper hemicontinuous compact-valued correspondence if for every sequence $\{y_n\}$ and every sequence $\{x_n\}$ with $y_n \in \Gamma(y_n)$, there is a convergent subsequence of $\{x_n\}$ with limit in $\Gamma(y)$.

Lower hemicontinuity for a compact-valued correspondence. A function $\Gamma : Y \to c(X)$ is a called a lower hemicontinuous compact-valued correspondence if for every sequence $_n \to y$ and every $x \in \Gamma(y)$ there exists a sequence $\{x_n\}$ such that $x_n \to x$ and $x_n \in \Gamma(y_n)\, \forall n$.

The following result says that Hausdorff continuity can be split up into u.h.c. and l.h.c.

Theorem. *A correspondence $Y \to c(X)$ is continuous iff it is l.h.c. and u.h.c.*

The following result is the closest one can come to a positive answer to the question above.

The Maximum Theorem. *Consider a family of optimization problems $(P_y), y \in Y$, of the following type: X and Y are metric spaces, $f : X \times Y \to \mathbb{R}$ is continuous, for each $y \in Y$ a nonempty compact set $\Gamma(y) \subseteq X$ is given, and assume that the correspondence Γ is continuous. For all $y \in Y$ we consider* *the problem*

$$f(x,y) \to \max, x \in \Gamma(y). \tag{P_y}$$

Then

1. *The solution correspondence G of the family $(P_y), y \in Y$ is compact-valued upper hemicontinuous,*
2. *The optimal value function V^* of the family $(P_y), y \in Y$ is continuous.*

This is what Acemoglu writes about this result: *"In this section, I state one of the most important theorems in mathematical economic analysis, Berge's Maximum Theorem. This theorem is not only essential for dynamic optimization, but it also plays a major role in general equilibrium theory, game theory, political economy, public finance, and industrial organization. In fact, it is hard to imagine any area of economics where it does not play a major role."*

Now we can explain how the concept Hausdorff continuity can be avoided (and this is what is usually done). You just **define** continuity of a correspondence as follows: the correspondence is u.h.c. and l.h.c. The reason to consider the maximum theorem in this course is not that it gives the next best thing to a positive answer to a natural and interesting question. The reason is that it is a useful tool to establish valuable results on the social planner problem.

Proof of the maximum theorem. Let $\{y_n\}$ be a sequence in Y with limit y. Choose $x_n \in G(y_n)$ as we may by Weierstrass. Choose a subsequence $\{x_{n_k}\}$ that converges to an element x of $\Gamma(y)$, using that Γ is u.h.c. For each $z \in \Gamma(y)$, we choose a sequence $z_k \to z$, with $z_k \in \Gamma(y_{n_k})\,\forall k$, using that Γ is l.s.c. Since $f(x_{n_k}, y_{n_k}) \geq f(z_k, y_{n_k})\,\forall k$, and f is continuous, we get $f(x,y) \geq f(z,y)$. As $z \in \Gamma(y)$ is arbitrary, we get $x \in G(y)$. Hence G is u.h.c.

Let $y_n \to y$. Choose $x_n \in G(y_n)\,\forall n$. Take a subsequence $\{x_{n_i}\}$ such that $f(x_{n_i}, y_{n_i})$ converges for $i \to \infty$. Choose a subsequence $\{x_{n_{i_j}}\}$ that converges to $x \in G(y)$, using that G is u.h.c. Then $f(x_{n_{i_j}}, y_{n_{i_j}}) \to f(x,y)$. As $f(x_{n_i}, y_{n_i})$ converges for $i \to \infty$, it has the same limit $f(x,y)$ as its subsequence $f(x_{n_{i_j}}, y_{n_{i_j}})$. Thus we have shown that all convergent subsequences of $f(x_n, y_n)$ have the same limit $f(x,y)$. Therefore, the sequence $f(x_n, y_n)$ converges with limit $f(x,y)$. We have $f(x,y) = V^*(y)$ as $y \in G(x)$ and we have for all n that $f(x_n, y_n) = V^*(y_n)$ as $y_n \in G(y_n)$. Thus we get that the sequence $V^*(y_n)$ converges with limit $V^*(y)$. Hence V^* is continuous.

5.3 Continuity optimal value function infinite horizon problem

Theorem (Continuity value function). *The optimal value function V^* of the infinite horizon problem is continuous and bounded. Moreover, there exists an optimal policy.*

Proof. We use the first one of the two tricks given above for applying the contraction mapping theorem. Consider the subset $C(X)$ of continuous bounded functions of the metric space $B(X)$ of bounded functions on X defined above. It is closed. Consider again the operator T given on functions $V : X \to \mathbb{R} \cup \{\pm\infty\}$ by

$$TV(x) = \sup_{y \in G(x)} U(x,y) + \beta V(y).$$

By Weierstrass' theorem, the right hand side minimization problem has a solution. The Maximum theorem gives that TV is continuous if $V \in C(X)$: indeed , we have here a family of optimization problems

$$U(x,y) + \beta V(y) \to \max, y \in G(x) \qquad (Q_x).$$

Note that compared to the Maximum Theorem, the roles of x and y are reversed here: here y is the quantity that has to be chosen in an optimal way ('the variable of optimization') and x is the parameter in the problem. Now we start with an arbitrary function $V \in C(X)$ and apply T repeatedly. In the limit we get the fixed point, by the contraction mapping theorem. This shows that the optimal value function V^* is continuous. It remains to prove the existence of an optimal policy: for this it suffices to apply Weierstrass to the right hand side of

$$TV^*(x) = \max_{y \in G(x)} (U(x,y) + \beta V^*(y)).$$

5.4 Strict concavity of the value function

We recall here that for a vector space X, a subset $A \subset X$ and a function $f : A \to \mathbb{R}$, the function f is called convex if its epigraph $\{(a,\rho) | a \in A, \rho \in \mathbb{R}, \rho \geq f(a)\}$ is a convex subset of the vector space $X \times \mathbb{R}$. A consequence of this definition is that for a convex function f its domain A is automatically a convex set; we do not display the verification here. Moreover, we recall that for a vector space X, a subset $A \subset X$ and a function $f : A \to \mathbb{R}$, the function f is called concave if minus the function, $-f$, is convex. Therefore, the domain of a concave function is a convex set.

The social planner problem is about maximizing total utility. It is natural to require that utility functions are concave. An exaggeration gives the intuition behind this: an additional dollar gives more additional pleasure to a poor man than to a rich man. That is the utility of your wealth has a derivative that is decreasing: this gives that the utility function is concave.

Theorem (Concavity value function). *Assume $U : X_G \to \mathbb{R}$ is strictly concave. Then the value function V^* is concave.*

Proof. Let $C'(X) = \{V \in C(X) | V \text{ is concave}\}$ and $C''(X) = \{V \in C(X) | V \text{ is strictly concave}\}$. The set $C'(X)$ is closed in $C(X)$, clearly, so as a metric space $C'(X)$ is complete, as $C(X)$ is complete. However $C''(X)$ is not closed in $C(X)$, so it is not complete. We use the second of the two tricks given above for applying the contraction mapping theorem. Therefore, to prove the theorem, it suffices to prove the implication $V \in C'(X) \Rightarrow TV \in C''(X)$. Assume $V \in C'(X)$. Choose $x_i \in X, i = 1, 2$ with $x_1 \neq x_2$ and $\alpha \in (0, 1)$. We pick $y_i \in G(x_i), i = 1, 2$ such that $TV(x_i) = U(x_i, y_i) + \beta V(y_i), i = 1, 2$, using the theorem of Weierstrass. Define $x_\alpha = (1-\alpha)x_1 + \alpha x_2$ and $y_\alpha = (1-\alpha)y_1 + \alpha y_2$. By assumption, the domain of $U : X_G \to \mathbb{R}$ is convex, and so it contains along with $(x_i, y_i), i = 1, 2$ also (x_α, y_α); that is $y_\alpha \in G(x_\alpha)$. Therefore, $TV(x_\alpha) \geq U(x_\alpha) + \beta V(y_\alpha)$. We have $U(x_\alpha) > (1-\alpha)U(x_1) + \alpha U(x_2)$ as U is strictly concave, and $V(y_\alpha) \geq (1-\alpha)U(y_1) + \alpha V(y_2$ as V is concave. In all, we get, on rearranging terms, $TV(x_\alpha) > (1-\alpha)[U(x_1, y_1) + \beta V(y_1)] + \alpha[U(x_2, y_2) + \beta V(y_2)] = (1-\alpha)TV(x_1) + \alpha TV(x_2)$. This proves that TV is strictly concave.

Corollary. *The problem (P) has a unique solution $\widehat{x}(\cdot)$. It can be obtained recursively as $\widehat{x}(t+1) = \pi(\widehat{x}(t)))$, where $\pi : X \to X$ is a continuous policy in closed loop form.*

Proof. The set $M(\xi)$ consists of precisely one element as it is the solution set of a strict concave solvable problem $U(x, y) + V^*(y) \to \max, x \in G(y)$. Therefore $M(x)$ is single-valued, that is, there is a unique policy function in closed loop form, and so, as $M(x)$ is upper hemicontinuous, π is continuous.

5.5 Strict monotonicity of the value function

Theorem (Monotonicity value function). *Assume that for each $y \in X$, $U(\cdot, y)$ is strictly increasing in each one of its arguments, and assume that $x \leq x'$ implies $G(x) \subseteq G(x')$. Then V is strictly increasing in each one of its arguments.*

Proof. Let $C'(X)$ be the set of bounded, continuous, nondecreasing functions on X, and let $C''(X)$ be the set of bounded, continuous, strictly increasing functions on X. $C'(X)$ is a closed subset of $C(X)$, and so $C'(X)$ is a complete metric space. However, $C''(X)$ is not closed and therefore $C'"(X)$ is not complete as a metric space. We use the second of the two tricks given above for applying the contraction mapping theorem. Consider any $V \in C'(X)$. By the assumption, the function $\max_{y \in G(x)}[U(x, y) + \beta V(y)]$ is strictly increasing. That is, $TV \in C''(X)$. This completes the proof.

5.6 Differentiability of the value function

Theorem (Differentiability value function). *Assume $U : X_G \to \mathbb{R}$ is strictly concave. and C^1 on the interior of its domain X_G. Let $c \in \mathrm{int}X$ and $g(c) \in \mathrm{int}G(c)$. Then V^* is continuously differentiable in c and $(V^*)'(c) = \frac{\partial U}{\partial x_i}[c, \pi(c)]$.*

To prove this result we need the following theorem.

Theorem. *Let V, W be concave functions on an open disk in $\mathbb{R}^n$ with center c for which $W \leq V$, $W(c) = V(c)$ and W differentiable at c. Then V is differentiable in c and $W'(c) = V'(c)$.*

Proof. Any subgradient p of V at c must satisfy $p(x-c) \geq V(x) - V(c) \geq W(x) - W(c)$ for all x in the disk. As W is differentiable, p is unique, and so V is a concave function with a unique subgradient. Therefore, V is differentiable in c.

Proof differentiability theorem. We have $g(c) \in \text{int}G(x)$ and G is continuous, so $\pi(c) \in \text{int}G(x)$ for all x in some neighborhood of c. Define W on this neighborhood by

$$W(x) = U(x, \pi(c)] + \beta V^*(g(c)).$$

Since U is concave and differentiable, W is concave and differentiable. Moreover, as $\pi(c) \in G(x)$ for all x in the neighborhood, it follows that

$$W(x) \leq \max_{y \in G(x)} [U(x,y) + \beta V^*(y)] = V^*(x)$$

for all x in the neighborhood, with equality at c. So by the theorem above, we get the desired results.

5.7 Problems on the lecture notes

1. Prove that for a contraction T with factor β and fixed point s one has $d(T^n\bar{s}, s) \leq \beta^n d(\bar{s}, s)$ for all $n \in \mathbb{N}$ and all starting points $\bar{s}$.

2. Prove the first statement in the last paragraph of section (5.1) ('real application ...').

3. Prove the second statement in the last paragraph of section (5.1) ('real application ...').

4. Prove the theorem of Weierstrass for metric spaces.

5. Determine the Hausdorff distance between A = unit disk = $\{(x,y)|x^2 + y^2 \leq 1\}$ and B = unit square = $\{(x,y)|\max(|x|,|y|) \leq 1\}$.

Your tasks between lecture 5 and lecture 6. You are expected to do as follows.

- In groups of two you make the five problems on the lecture notes given above, as well as the following four problems from Introduction to Modern Economic Growth by Acemoglu (page 224-225). You have to hand in your work on time. Present complete solutions.

  1. 6.6
  2. 6.7
  3. 6.8

4. 6.9

- Study these lecture notes carefully
- Try to look through chapter 6 section 6.1-6.3, 6.6 till p.207 (example 6.5 is not included) and Appendix A.1-A.6 of Introduction to Modern Economic Growth by Acemoglu. Moreover study the notes of Bjoern Bruegemann on the same material.

  The aim of the notes of this lecture and the last part of the notes of the previous lecture is to give an account of this material. I have also profited from the great notes by Bjoern Bruegemann.