Version: 10-10-2011

2.1 Optimality conditions in unconstrained optimization

Recall the definitions of *global, local minimizer*.

Example: the ‘humpback function’

$$\min \; x_1^2(4 - 2.1x_1^2 + \frac{1}{3}x_1^4) + x_1x_2 + x_2^2(-4 + 4x_2^2).$$

Two global minima: (0.089 -0.717) and (-0.0898 0.717), and four strict local minima.

Fermat's theorem, Necessary condition

Th. 2.5 (*Fermat*) Let $f : \mathbb{R}^n \to \mathbb{R}$ be in C^1 (C^2). If the point $\overline{x} \in \mathbb{R}^n$ is a local minimizer of the function f then $\nabla f(\overline{x}) = 0$ (and $\nabla^2 f(\overline{x})$ is psd).

Ex. The conditions $\nabla f(\overline{x}) = 0$, $\nabla^2 f(\overline{x})$ psd, are necessary but not sufficient minimality conditions.
Take $f(x) = x^3$ or $\pm x^4$ as examples.

Sufficient conditions for local minimizer

Cor. 2.8 Let $f : \mathbb{R}^n \to \mathbb{R}$ be in C^2. If $\nabla f(\overline{x}) = 0$ and $\nabla^2 f(\overline{x})$ is positive definite then the point $\overline{x}$ is a strict local minimizer of the function f.

Global minimizer of convex functions

For convex functions *the situation is "easier"*

L2.4 *(Ex2.1)* **Any (strict) local minimizer of a convex function f is a (strict) global minimizer.**

T.2.6 **Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and in C^1. The point $\overline{x} \in \mathbb{R}^n$ is a global minimizer of f if and only if $\nabla f(\overline{x}) = 0$.**

Ex. (convex quadratic function) *With positive definite Q, $c \in \mathbb{R}^n$ consider the quadratic function*

$$f(x) := \frac{1}{2} x^T Q x + c^T x.$$

Now, f is convex and therefore $\overline{x}$ is a global minimizer iff

$$\nabla f(\overline{x}) = 0 \quad \text{or} \quad Q\overline{x} + c = 0$$

The global minimizer is therefore given by $\overline{x} = -Q^{-1}c$.

2.2 Optimality conditions for constrained (convex) optimization

We consider the convex optimization problem:

$$(CO) \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_j(x) \le 0, \quad j \in J \\ & x \in \mathcal{C}, \end{array}$$

where $J = \{1, \cdots, m\}$,

- $\mathcal{C} \subseteq \mathbb{R}^n$ is a *convex (closed) set*;
- $f, g_1, \cdots, g_m$ are convex functions on $\mathcal{C}$ (or on an open set containing $\mathcal{C}$).
- The set of feasible points (feasible set) is:

$$\mathcal{F} = \{x \in \mathcal{C} \mid g_j(x) \le 0, \quad j = 1, \cdots, m\}.$$

Ex. Show that $\mathcal{F}$ is a convex set.

Consider first

$$P_0 : \quad \min\{ f(x) \ : \ x \in \mathcal{C}\},$$

with $\mathcal{C}$, a *relatively open convex* set and f is convex, differentiable on $\mathcal{C}$.

Th.2.9 The point $\overline{x} \in \mathcal{C}$ is a (global) *minimizer* of P_0 if and only if

$$\nabla f(\overline{x})^T s = 0 \text{ for all } s \in \mathcal{L},$$

where $\mathcal{L}$ denotes the *linear subspace* defined by $\mathrm{aff}(\mathcal{C}) = x + \mathcal{L}$ for any $x \in \mathcal{C}$.
(Here $\mathrm{aff}(\mathcal{C})$ is the affine hull of $\mathcal{C}$.)

Remark: *Essentially, P_0 is an "unconstrained problem"*

Example 1 Consider the relative interior of the simplex $S = \{x \in \mathbb{R}^3 \mid \sum_{i=1}^{3} x_i = 1,\ x \geq 0\}$ in $\mathbb{R}^3$:

$$C = \{x \in \mathbb{R}^3 \ : \ \sum_{i=1}^{3} x_i = 1,\ x > 0\},$$

and define the (convex!) function $f : \mathcal{C} \mapsto \mathbb{R}$ by

$$f(x) = (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2.$$

By Exercise 1.6*: aff*$(\mathcal{C}) = (1,0,0) + \mathcal{L}$, *where* $\mathcal{L}$ *is the* linear subspace

$$\mathcal{L} := \{s \in \mathbb{R}^3 \ : \ s_1 + s_2 + s_3 = 0\} = \alpha \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$\alpha, \beta \in \mathbb{R}$. *Now* $\overline{x} \in \mathcal{C}$ *is a* minimizer *of* $\min_{x \in \mathcal{C}} f(x)$, *if and only if* $\nabla f(\overline{x})^T s = 0$ *for all* $s \in \mathcal{L}$.

Since

$$\nabla f(x) = 2 \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \\ x_3 - 1 \end{bmatrix}$$

this holds if and only if $x_1 - 1 = x_2 - 1$ *and* $x_1 - 1 = x_3 - 1$, *which gives* $x_1 = x_2 = x_3$. *Since* $x_1 + x_2 + x_3 = 1$, *it follows that* f *has a unique minimizer, namely*

$$\bar{x} = (1/3; 1/3; 1/3).$$

Feasible Directions *The next characterization of minimizers is based on the concept of "feasible directions"*

Def.2.10 The vector $s \in \mathbb{R}^n$ is a *feasible direction* at a point $x \in \mathcal{F}$ if there is a $\lambda_0 > 0$ such that

$$x + \lambda s \in \mathcal{F} \quad \text{for all} \quad 0 \le \lambda \le \lambda_0 .$$

The *set of feasible directions* at $x \in \mathcal{F}$ is denoted by $\mathcal{FD}(x)$. Note that $0 \in \mathcal{FD}(x)$.

Recall. $K \subseteq \mathbb{R}^n$ is a *cone* if: $x \in K, \lambda > 0 \Rightarrow \lambda x \in K$.

L.2.13 For any $x \in \mathcal{F}$ the set $\mathcal{FD}(x)$ is a *convex cone*.

Th.2.14 (Let $f \in C^1$.) The feasible point $\overline{x} \in \mathcal{F}$ is a (global) *minimizer* of (CO) if and only if

$$\delta f(\overline{x}, s) = \nabla f(\overline{x})^T s \ge 0, \quad \forall s \in \mathcal{FD}(\overline{x}).$$

Example 1. *(continued): Reconsider the simplex in* $\mathbb{R}^3$*:*

$$\mathcal{C} = \{x \in \mathbb{R}^3 \ : \ \sum_{i=1}^{3} x_i = 1, \ x \geq 0\},$$

and $f : \mathcal{C} \mapsto \mathbb{R}$ *given by*

$$f(x) = (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2.$$

At the point $\overline{x} = (1/3, 1/3, 1/3)$*, the cone of feasible directions is simply*

$$\mathcal{FD}(\overline{x}) = \{s \in \mathbb{R}^3 \ : \ s_1 + s_2 + s_3 = 0\}. \text{ (why?)}$$

Therefore, $\overline{x} = (1/3, 1/3, 1/3)$ *is an* optimal solution *of* $\min_{x \in \mathcal{C}} f(x)$*, because*

$$\nabla f(\overline{x}) = 2 \begin{bmatrix} \overline{x}_1 - 1 \\ \overline{x}_2 - 1 \\ \overline{x}_3 - 1 \end{bmatrix} = -\frac{4}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

and hence $\nabla f(\overline{x})^T s = 0$ *for all* $s \in \mathcal{FD}(\overline{x})$.

Optimality conditions based on the Karush-Kuhn-Tucker relations: ***The conditions in Theorem 2.9/2.14 are not easy to check (in practice). We are thus interested in 'easier' minimality conditions.***

Definition 2.15

With respect to (CO), an index $j \in J$ is called *active* in $\overline{x} \in \mathcal{F}$ if $g_j(\overline{x}) = 0$.

The set of all active indices, also called *active index set* in $\overline{x}$, is denoted by $J_{\overline{x}}$.

Let us assume: $\mathcal{C} = \mathbb{R}^n$, $f, g_j \in C^1$ (*in the rest of this subsection*)

Def. The feasible point $\overline{x} \in \mathcal{F}$ satisfies the ***Karush-Kuhn-Tucker*** (KKT) conditions if there exist ***multipliers*** $\overline{y}_j$, $j \in J$ such that:
$\overline{y}_j \geq 0$, $\forall j \in J$ and

$$\nabla f(\overline{x}) = -\sum_{j \in J} \overline{y}_j \nabla g_j(\overline{x}), \quad \overline{y}_j g_j(\overline{x}) = 0, \ \forall j \in J \, .$$

The KKT-conditions can equivalently be given as

$$\nabla f(\overline{x}) = -\sum_{j \in J_{\overline{x}}} \overline{y}_j \nabla g_j(\overline{x}), \quad \overline{y}_j \geq 0, \quad \forall \, j \in J_{\overline{x}}.$$

Th.2.16 (*Sufficient condition*) If $\overline{x} \in \mathcal{F}$ satisfies the KKT conditions then $\overline{x}$ is a minimizer of (CO) (with $\mathcal{C} = \mathbb{R}^n$).

<u>Ex.</u> The KKT conditions are not necessary for minimality.
Take as example:

$$\min\{x \;:\; x^2 \leq 0\}.$$

The minimum occurs at $x = 0$ *(which is the only feasible point!). But there is no* y *such that*

$$1 = -y \cdot 0, \quad y \cdot 0 = 0, \quad y \geq 0.$$

(The problem here is that the feasible set $\mathcal{F} = \{0\}$ *does not have interior points.)*

Some examples: (*Application of the KKT-conditions*)

Ex. (minimizing a linear function over a sphere)

Let $a, c \in \mathbb{R}^n$, $c \neq 0$. Consider the problem

$$\min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^n c_i x_i \;\mid\; \sum_{i=1}^n (x_i - a_i)^2 \leq 1 \right\}.$$

The objective function is linear (hence convex), and there is one constraint function $g_1(x)$ ($m = 1$), which is convex. The KKT conditions give

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = -2\lambda \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix},$$

$$\lambda\Big(\sum_{i=1}^n (x_i - a_i)^2 - 1\Big) = 0, \quad \lambda \geq 0.$$

$c \neq 0$ implies $\lambda > 0$ and $\sum_{i=1}^{n}(x_i - a_i)^2 = 1$. We conclude that $x - a = -\alpha c$, for $\alpha = 1/(2\lambda) > 0$. Since $\|x - a\| = 1$, we have $\alpha \|c\| = 1$. So, the minimizing vector is

$$x = a - \alpha c = a - \frac{c}{\|c\|}.$$

Hence, the minimal value of the objective function is:

$$c^T x = c^T a - \frac{c^T c}{\|c\|} = c^T a - \frac{\|c\|^2}{\|c\|} = c^T a - \|c\|.$$

Ex. (Exercise 2.7)

Find a cylindrical can with height h, radius r and volume at least V such that the total surface area is minimal. Denoting the surface by $p(h,r)$, we want to solve

$$\min \ p(h,r) := 2\pi(r^2 + rh) \quad \text{s.t.} \quad \pi r^2 h \geq V, \ (r, h > 0)$$

This is not a convex optimization problem! (why?) Setting $r = e^{x_1}$ *and* $h = e^{x_2}$ *the problem becomes:*

$$\min \left\{ 2\pi \left(e^{2x_1} + e^{x_1+x_2} \right) \mid \pi \left(e^{2x_1+x_2} \right) \geq V \right\}$$

which is equivalent to

$$\min\{2\pi \left(e^{2x_1} + e^{x_1+x_2} \right) \mid \ln \frac{V}{\pi} - 2x_1 - x_2 \leq 0,$$

$$x_1 \in \mathbb{R}, \ x_2 \in \mathbb{R}\}.$$

This is a convex optimization problem! (why?)

Ex. Exercise 2.7 (ctd) *Let*

$$f(x) = 2\pi\left(e^{2x_1} + e^{x_1+x_2}\right) \text{ and } g(x) = \ln\frac{V}{\pi} - 2x_1 - x_2.$$

Now the problem becomes:

$$\min\left\{f(x) \mid g(x) \le 0,\ x \in \mathbb{R}^2\right\}.$$

KKT conditions for this problem: we are looking for some $\bar{x} \in \mathcal{F}$ *such that*

$$\nabla f(\bar{x}) = -\bar{y}\nabla(g(\bar{x}))$$

for some scalar $\bar{y} \ge 0$, *such that* $\bar{y}g(\bar{x}) = 0$. *The KKT conditions are:*

$$\nabla f(\bar{x}) = -\bar{y}\nabla g(\bar{x}),\ \bar{y}g(\bar{x}) = 0,\ \bar{y} \ge 0.$$

Ex. Exercise 2.7 (ctd.) *Calculate the gradients to get:*

$$(\star) \quad 2\pi \begin{bmatrix} 2e^{2\overline{x}_1} + e^{\overline{x}_1+\overline{x}_2} \\ e^{x_1+x_2} \end{bmatrix} = -\overline{y} \begin{bmatrix} -2 \\ -1 \end{bmatrix} = \overline{y} \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

and $\overline{y}g(\overline{x}) = 0$, $\overline{y} \geq 0$. *From* $(\star)$ *we derive:* $\overline{y} > 0$ *and therefore*

$$\frac{2e^{2\overline{x}_1} + e^{\overline{x}_1+\overline{x}_2}}{e^{\overline{x}_1+\overline{x}_2}} = \frac{2\overline{y}}{\overline{y}} = 2.$$

This implies $2e^{\overline{x}_1} = e^{\overline{x}_2}$, *whence* $2r = h$.

Conclusion: **The surface is minimal if $2r = h$.**
Now use $g(\overline{x}) = 0$, to get r and h in terms of V.

Remark *Later (in Chapter 6 "Nonlinear Optimization") we will be able to solve the problem without the "transformation trick".*

Question to be answered now: *Under which condition is the KKT-relation necessary for minimality?*

2.2.2 The Slater condition

Def.2.18 $x^0 \in \mathcal{C}^0$ is called a ***Slater point*** of (CO) if

$$\begin{array}{llll} g_j(x^0) < 0, & \forall j \text{ where} & g_j \text{ is nonlinear,} \\ g_j(x^0) \le 0, & \forall j \text{ where} & g_j \text{ is (affine) linear.} \end{array}$$

We say that (CO) satisfies the ***Slater condition*** (constraint qualification) if a Slater point exists.

Example. (*Slater regularity:*)
Consider the optimization problem

$$\begin{array}{rl} \textbf{min} & f(x) \\ \textbf{s.t.} & x_1^2 + x_2^2 \leq 4 \\ & x_1 - x_2 \geq 2 \\ & x_2 \geq -1 \\ & \mathcal{C} = \mathbb{R}^2. \end{array}$$

The points $(1, -1)$ **and** $(\frac{3}{2}, -\frac{3}{4})$ **are Slater points,** $(2, 0)$ **not.**

Definition **Some constraints $g_j(x) \leq 0$ may take the value zero at all feasible points.** *Each such constraint can be replaced by the equality constraint $g_j(x) = 0$, without changing the feasible region.* **These constraints are called *singular* while the others are called *regular*.**

Under the Slater condition, we define the index sets J_s of *singular-* and J_r of *regular constraints*:

$$J_s := \{j \in J \mid g_j(x) = 0, \ \forall x \in \mathcal{F}\},$$

$$J_r := J \setminus J_s = \{j \in J \mid g_j(x) < 0 \text{ for some } x \in \mathcal{F}\}.$$

Remark: *Note, that if (CO) satisfies the Slater condition, then all singular constraints must be linear.*

Ideal Slater points

Def.2.20 A Slater point $x^* \in \mathcal{C}^0$ is called an ***Ideal Slater point*** of (CO) if

$$\begin{array}{ll} g_j(x^*) < 0 & \text{for all } j \in J_r, \\ g_j(x^*) = 0 & \text{for all } j \in J_s. \end{array}$$

L.2.21 If (CO) has a Slater point then it also has an ideal Slater point $x^* \in \mathcal{F}$.

Rem: An ideal Slater point is in the relative interior of $\mathcal{F}$ (Exercise 2.10).

Example. (*Slater regularity, continued:*) **Reconsider the program**

$$\begin{array}{ll} \textbf{min} & f(x) \\ \textbf{s.t.} & x_1^2 + x_2^2 \le 4 \\ & x_1 - x_2 \ge 2 \\ & x_2 \ge -1 \\ & \mathcal{C} = \mathbb{R}^2. \end{array}$$

The point $(1, -1)$ is a Slater point, but not an ideal Slater point.

The point $(\frac{3}{2}, -\frac{3}{4})$ is an ideal Slater point.

2.2.3 Convex Farkas Lemma

Remark ***The so-called convex Farkas Lemma is the basis for the duality theory in convex optimization*** **We begin with an auxiliary result.**

Th.2.23 *(see also [FKS,Th10.1)* **Let** $\emptyset \neq U \subset \mathbb{R}^n$ **be convex and** $w \notin U$**. Then there exists a separating hyperplain** $H = \{x \mid a^T x = \alpha\}, 0 \neq a \in \mathbb{R}^n, \alpha \in \mathbb{R}$ **such that**

$$a^T w \geq \alpha \geq a^T x \quad \forall x \in U$$

and $\alpha > a^T u_0$ **for some** $u_0 \in U$**.**

Proof: See pdf-file "extraproofs" for a proof.

Observation: A number a is a lower bound for the optimal value of (CO) if and only if the system

$$(1) \qquad \begin{array}{ll} f(x) < a & \\ g_j(x) \leq 0, & j = 1, \ldots, m \\ x \in \mathcal{C} & \end{array}$$

has no solution.

Lemma.2.25 (Farkas) Let (CO) satisfy the Slater condition. Then the inequality system (1) has no solution if and only there exists a vector $y = (y_1, \cdots, y_m)$ such that

$$(2) \qquad y \geq 0, \quad f(x) + \sum_{j=1}^{m} y_j g_j(x) \geq a \, , \; \forall x \in \mathcal{C}.$$

The systems (1) and (2) are called *alternative systems*, because exactly one of them has a solution.

Proof of the (convex) Farkas lemma

We need to show that precisely one of the systems (1) or (2) has a solution.

- **First we show (easy part) that not both systems can have a solution:**
  Otherwise for some $x \in \mathcal{C}$ and $y \geq 0$, the relation

  $$a \leq f(x) + \sum_{j=1}^{m} y_j g_j(x) \leq f(x) < a,$$

  would hold, a contradiction.
- **Then we show that at least one of the two systems has a solution.**
  To do so, we show that if (1) has no solution. then (2) has a solution.

Proof: Define the set $\mathcal{U} \subset \mathbb{R}^{m+1}$ as follows.

$$\mathcal{U} = \{u = (u_0; \ldots ; u_m) \mid \exists x \in \mathcal{C} \ \text{ such that}$$

$$\left\{ \begin{array}{rcl} f(x) & < & a + u_0 \\ g_j(x) & \leq & u_j \ \text{ if } j \in J_r \\ g_j(x) & = & u_j \ \text{ if } j \in J_s \end{array} \right.$$

Since (1) is infeasible, $0 \notin \mathcal{U}$. One easily checks that $\mathcal{U}$ is a nonempty convex set (using that the singular functions are linear). So there exists a hyperplane that separates 0 from $\mathcal{U}$, i.e., there exists a nonzero vector $y = (y_0, y_1, \cdots, y_m)$ such that

$$\begin{array}{rcl} y^T u & \geq & 0 = y^T 0 \quad \text{for all } u \in \mathcal{U} \\ y^T \overline{u} & > & 0 \quad \text{for some } \overline{u} \in \mathcal{U}. \end{array} \qquad (*)$$

The rest of the proof is divided into four parts.

I. Prove that $y_0 \geq 0$ and $y_j \geq 0$ for all $j \in J_r$.

II. Establish that

$$y_0(f(x) - a) + \sum_{j=1}^{m} y_j g_j(x) \geq 0 \;, \; \forall x \in \mathcal{C}.$$

holds

III. Prove that y_0 must be positive.

IV. Show by induction that we can assume $y_j > 0, \; \forall j \in J_s$.

Example. *(Illustration of the Farkas lemma)*

Let us consider the convex optimization problem

$$(CO) \quad \min \quad 1 + x$$
$$\text{s.t.} \quad x^2 - 1 \leq 0$$
$$x \in \mathbb{R}.$$

Then (CO) is Slater regular. The optimal value is 0 (for $x = -1$). Hence, the system

$$1 + x < 0, \quad x^2 - 1 \leq 0, \quad x \in \mathbb{R}$$

has no solution. By the Farkas lemma there must exist a real number $y \geq 0$ such that

$$1 + x + y\left(x^2 - 1\right) \geq 0, \qquad \forall x \in \mathbb{R}.$$

Indeed, taking $y = \frac{1}{2}$ we get

$$x + 1 + y(x^2 - 1) = \frac{1}{2}x^2 + x + \frac{1}{2} = \frac{1}{2}(x + 1)^2 \geq 0.$$

Exercise 2.11 (*"Linear Farkas Lemma"*) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Exactly one of the following alternative systems (I) or (II) is solvable:

$$(I) \qquad A^T x \geq 0, \qquad x \geq 0, \qquad b^T x < 0,$$

or

$$(II) \qquad Ay \leq b, \qquad y \geq 0.$$

Proof: *Apply Farkas' lemma to the special case:* $f(x) = b^T x$,

$$g_j(x) = -(A^T x)_j = -\left(a^j\right)^T x, \quad j = 1, \cdots, m$$

where a^j *denotes column* j *of* A, *and* $\mathcal{C}$ *is the nonnegative orthant (nonnegative vectors) of* $\mathbb{R}^m$. **Make use of:**

$$\mathbf{z^T x} = \sum_{\mathbf{i}} \mathbf{z_i x_i} \geq \mathbf{0} \ \forall \mathbf{x} \geq \mathbf{0} \Leftrightarrow \mathbf{z} \geq \mathbf{0} \qquad (\star)$$

2.2.4. Karush-Kuhn-Tucker theory (*Duality theory*)

Remark: This (Lagrange) Duality Theory is based on the concept of saddle points.

Def. We introduce the Lagrangean function of (CO):

$$L(x, y) := f(x) + \sum_{j=1}^{m} y_j g_j(x)$$

Note that $L(x,y)$ *is* convex *in* x *and* linear *in* y.

Note: For any $y \geq 0$:

$$\inf_{x \in \mathcal{C}} \{f(x) + \sum_{j=1}^{m} y_j g_j(x)\} \leq \inf_{\substack{x \in \mathcal{C} \\ g_j(x) \leq 0, \forall j}} \{f(x) + \overbrace{\sum_{j=1}^{m} y_j g_j(x)}^{\leq 0}\}$$

$$\leq \inf_{x \in \mathcal{C}} \{f(x) \mid g_j(x) \leq 0, \ \forall j\}$$

So: For any $y \geq 0$,

$$\inf_{x \in \mathcal{C}} \{f(x) + \sum_{j=1}^{m} y_j g_j(x)\} \leq \inf_{x \in \mathcal{C}} \{f(x) \mid g_j(x) \leq 0,\ \forall j\} \equiv \text{(CO)}$$

Note further that for any fixed x:

$$\sup_{y \geq 0} \{f(x) + \sum_{j=1}^{m} y_j g_j(x)\} = \begin{cases} f(x) & \text{if } g_j(x) \leq 0, \forall j \\ \infty & \text{otherwise} \end{cases}$$

By taking (above) the sup over $y \geq 0$ we find:

Weak Duality:

$$\sup_{y \geq 0} \inf_{x \in \mathcal{C}} L(x, y) \leq \inf_{x \in \mathcal{C}} \sup_{y \geq 0} L(x, y) \qquad (*)$$

The righthand side is "equivalent" with (CO);
the lefthand side is called the (Lagrangean) dual problem of (CO) (*See also Chapter 3*)

Def. 2.26 A vector pair $(\overline{x}, \overline{y})$, with $\overline{x} \in \mathcal{C}$ and $\overline{y} \geq 0$ is called a *saddle point* of the Lagrange function $L(x, y)$ if

$$L(\overline{x}, y) \leq L(\overline{x}, \overline{y}) \leq L(x, \overline{y}), \ \forall x \in \mathcal{C}, \ \forall y \geq 0.$$

Lemma 2.28: A saddlepoint $(\overline{x}, \overline{y})$ of $L(x, y)$ satisfies the strong duality relation

$$\sup_{y \geq 0} \inf_{x \in \mathcal{C}} L(x, y) = L(\overline{x}, \overline{y}) = \inf_{x \in \mathcal{C}} \sup_{y \geq 0} L(x, y) \ .$$

Theorem 2.30 Let the problem (CO) satisfy the Slater condition. Then the vector $\overline{x}$ is a minimizer of (CO) if and only if there is a vector $\overline{y}$ such that $(\overline{x}, \overline{y})$ is a saddle point of the Lagrange function L.

We are now able to prove the converse of Th.2.16.

Cor.2.33 Let $\mathcal{C} = \mathbb{R}^n$ and let $f, g_j \in C^1$ be convex functions. Suppose (CO) satisfies the Slater condition. Then $\overline{x}$ is a minimizer of (CO) if and only if with some $\overline{y} \geq 0$ the pair $(\overline{x}, \overline{y})$ satisfies the KKT condition.

Ex. (*without Slater condition*) Show that the pair $(\overline{x}, \overline{y})$ $(\overline{x} \in \mathcal{F}, \overline{y} \geq 0)$ satisfies the KKT condition if and only if $(\overline{x}, \overline{y})$ $(\overline{y} \geq 0)$ is a saddle point of L.