Convex optimization

Optimality conditions

Consider the primal-dual pair of problems with **differentiable** functions $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$:

$$\left.\begin{array}{lll} Min & f_0(x) & \\ & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}\right\} \quad (P), \qquad \left.\begin{array}{ll} Max & G(u,v) \\ & u \geq 0 \end{array}\right\} \quad (D),$$

associated with the Lagrangian

$$L(x,u,v) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x) + \sum_{i=1}^{p} v_i h_i(x),$$

and the dual function is defined as

$$G(u,v) = \inf_x L(x,u,v).$$

KKT (Karush-Kuhn-Tucker) optimality conditions:

$$f_i(x^*) \leq 0, \quad i = 1, \ldots, m,$$

$$h_i(x^*) = 0, \quad i = 1, \ldots, p,$$

$$u^* \geq 0, \quad i = 1, \ldots, m,$$

$$u_i f_i(x^*) = 0, \quad i = 1, \ldots, m$$

$$\nabla f_0(x^*) + \sum_{i=1}^m u_i^* \nabla f_i(x^*) + \sum_{i=1}^p v_i^* \nabla h_i(x^*) = 0.$$

If x^* is optimal for (P) and (u^*, v^*) is optimal for (D) and strong duality property $p^* = d^*$ holds, then

$$f_0(x^*) = L(x^*, u^*, v^*) = G(u^*, v^*).$$

Implications:

- Complementarity - $u_i f_i(x^*) = 0, \; i = 1, \ldots, m$
- If the functions $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ are differentiable, then x^* minimizes $L(x, u^*, v^*)$ and hence $\nabla_x L(x^*, u^*, v^*) = 0$.

For any problem with strong duality and differentiable functions it holds: If x^* and (u^*, v^*) are optimal solutions of (P) and (D), respectively, then they satisfy the system of KKT optimality conditions.

For **convex problems** with differentiable functions it holds:
If x^* and (u^*, v^*) satisfy the system of KKT conditions, then x^* is optimal for (P) and (u^*, v^*) is optimal for (D).

If we assume

- the problem is convex,
- the functions $f_0, f_1, \ldots, f_m, h_1, \ldots, h_p$ are differentiable
- Slater condition is satisfied

then the KKT conditions are necessary and sufficient conditions of optimality.

Importance:

- In some cases it is possible to solve the system of KKT conditions analytically to obtain the optimal solution.
- Many convex optimization algorithms are based on solving the system of KKT conditions.

Example: "Water filling"

Consider the convex optimization problem:

$$\left.\begin{array}{ll} Min & -\sum_{i=1}^{n} \ln(\alpha_i + x_i) \\ & \mathbf{1}^T x = 1, \\ & x \geq 0, \end{array}\right\} \quad (WF)$$

where $\alpha_i > 0$ are given.

- n communication channels
- x_i - transmitter power allocated to the i-th channel
- $\ln(\alpha_i + x_i)$ communication rate of the channel
- Problem: to allocate a total power of the channels in order to maximize the total communication rate

Lagrangian:

$$L(x,u,v) = -\sum_{i=1}^{n} \ln(\alpha_i + x_i) - u^T x + v(\mathbf{1}^T x - 1)$$

KKT conditions:

- Feasibility: $\mathbf{1}^T x = 1$, $x \geq 0$, $u \geq 0$;
- Complementarity: $u_i x_i = 0, \; i = 1, \ldots, m$;
- $\nabla_x L(x,u,v) = -\frac{1}{\alpha_i + x_i} - u_i + v = 0$

The variable u can be eliminated: $u = v - \frac{1}{\alpha_i + x_i}$

- Feasibility: $\mathbf{1}^T x = 1$, $x \geq 0$, $v \geq \frac{1}{\alpha_i + x_i}$;
- Complementarity: $x_i \left(v - \frac{1}{\alpha_i + x_i} \right) = 0, \; i = 1, \ldots, m$;

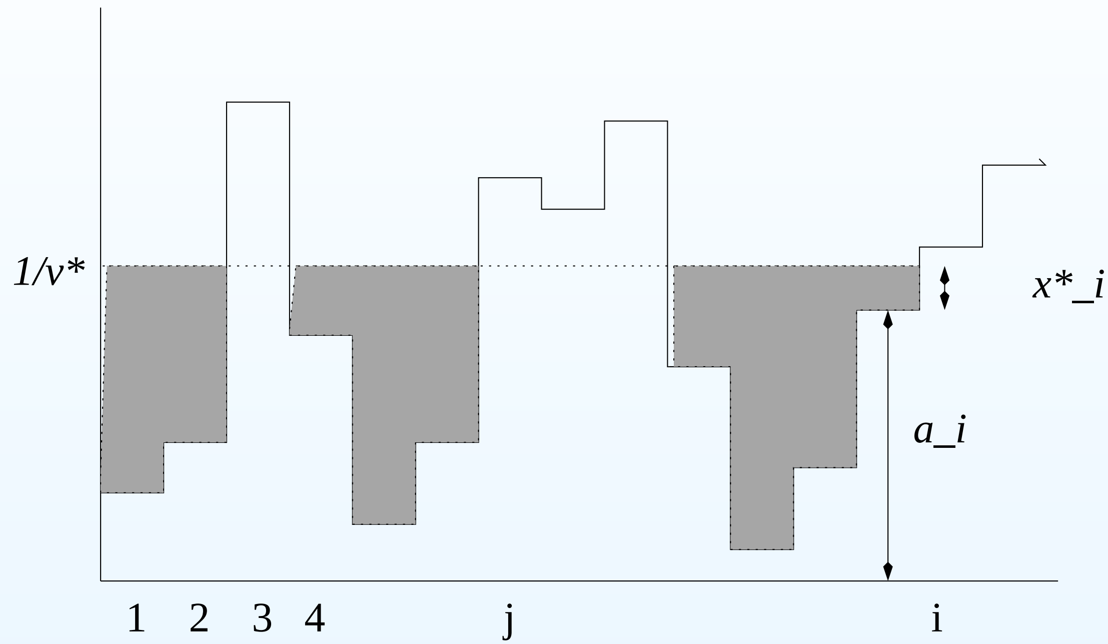
- If $v < \frac{1}{\alpha_i}$ then $x_i = \frac{1}{v} - \alpha_i$.
- If $v \geq \frac{1}{\alpha_i}$, then $x_i = 0$. Therefore

$$x_i = \max\left\{0, \frac{1}{v} - \alpha_i\right\}, \quad i = 1, \ldots, m$$

a

$$\sum_{i=1}^{n} \max\left\{0, \frac{1}{v} - \alpha_i\right\} = 1.$$

The function on the left is piece-wise linear and increasing in $\frac{1}{v}$ so the equation has a unique solution.



"Water filling" method: The height of each patch is given by α_i. The region is flooded to a level $1/v^*$ which uses a total quantity of water equal to one. The height of the water above each patch is the optimal value of x_i^*.

PERTURBATION AND SENSITIVITY ANALYSIS

$$\left.\begin{array}{lll} Min & f_0(x) & \\ & f_i(x) \leq r_i, & i = 1, \ldots, m, \\ & h_i(x) = s_i, & i = 1, \ldots, p. \end{array}\right\} \quad (P(r,s))$$

- $r_i > 0$ - we loosen the i-th constraint
- $r_i < 0$ - we tighten the i-th constraint
- Optimal value of the perturbed problem (P(r,s)):

$$p^*(r,s) = \inf\{f_0(x) \mid f_i(x) \leq r_i, i = 1, \ldots, m,\ h_i(x) = s_i, i = 1, \ldots, p\}$$

- $p^*(0,0) = p^*$
- If (P) is a convex optimization problem, then $p^*(r,s)$ is a convex function of (r,s)

Assume that strong duality holds $d^* = p^*$ and the dual optimum is attained. Let (u^*, v^*) be the optimal solution of the dual problem. Then:

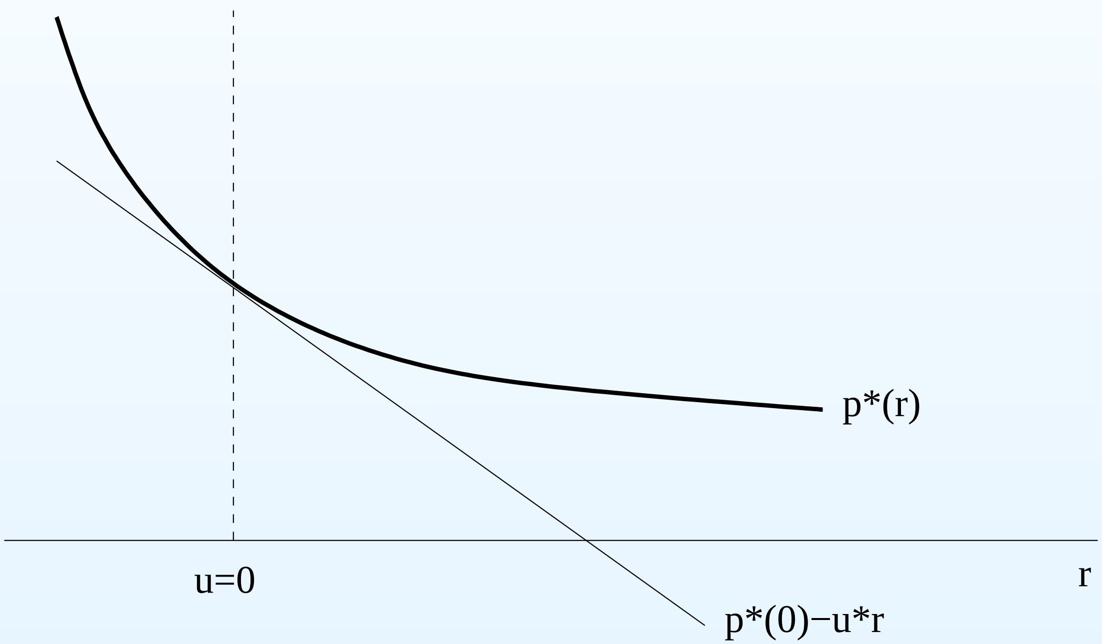
$$p^*(r, s) \geq p^*(0, 0) - (u^*)^T r - (v^*)^T s$$

Conclusions:

- If u_i^* is large and we tighten the i-th constraint (choose $r_i < 0$), then the optimal value $p^*(r, s)$ increases greatly.
- If v_i^* is large and positive and we take $s_i < 0$ or if v_i^* is large and negative and we take $s_i > 0$, then the value $p^*(r, s)$ increases greatly.
- If u_i^* is small and we loosen the i-th constraint (choose $r_i > 0$), then the value $p^*(r, s)$ will not decrease much.
- If v_i^* is small and positive and we take $s_i > 0$ or if v_i^* is small and negative and we take $s_i < 0$, then the value $p^*(r, s)$ will not decrease much.

Topic 6: Optimality conditions

$$\left.\begin{array}{ll} Min & f_0(x) \\ & f_1(x) \leq r \end{array}\right\}$$



The optimal value $p^*(r)$ of the perturbed convex problem with one constraint $f_1(x) \leq r$is a convex function of r. Affine function $p^*(0) - u^*r$ is the lower bound on $p^*(r)$.

Local sensitivity analysis

- Assume $p^*(r, s)$ is differentiable at $r = 0, s = 0$.
- Let $r = te_i,\ s = 0$. Then

$$\lim_{t\to 0} \frac{p^*(te_i,0)-p^*(0,0)}{t} = \frac{\partial p^*(0,0)}{\partial r_i}$$

- For $t > 0$ we have

$$\frac{p^*(te_i,0)-p^*(0,0)}{t} \geq -u_i^*$$

- For $t < 0$ we have

$$\frac{p^*(te_i,0)-p^*(0,0)}{t} \leq -u_i^*$$

- Therefore

$$\frac{\partial p^*(0,0)}{\partial r_i} = -u_i^*$$

- Analogously we obtain

$$\frac{\partial p^*(0,0)}{\partial s_i} = -v_i^*$$

Hence it holds:

$$\frac{\partial p^*(0,0)}{\partial r_i} = -u_i^*, \quad \frac{\partial p^*(0,0)}{\partial s_i} = -v_i^*$$

If x^* is optimal for (P):

- $f_i(x^*) < 0$ - inactive constraint - $u_i^* = 0$
- $f_i(x^*) = 0$ - active constraint-
  - u_i^* is small - the constraint can be loosened/tightened without much effect on the optimal value p^*.
  - u_i^* is large - if the constraint is loosened/ tightened a bit, the effect on the optimal value will be great