OPT/SMF (b)

L05

Pokročilé metody: neúplná měření, konvexní programování

informačně neúplná měření

princip maximální entropie (ME)

- volím řešení s maximální entropií
- vede na nejméně informativní odhad

QT – von Neumannova entropie

$\boldsymbol{E}(\rho) = -\mathrm{Tr}(\rho \log \rho)$ ... zobecnění Shannonovy entropie

vazby

$$\mathrm{Tr}(\rho \Pi_j) = \boldsymbol{p}_j$$

hledám $\max \boldsymbol{E}(\rho)$ s vazbami → multiplikátory

$$\max \; -\mathrm{Tr}(\rho \log \rho) + \sum_j \lambda_j \mathrm{Tr}(\rho \Pi_j)$$

$$\hat{\rho}_{\mathrm{ME}} \propto \boldsymbol{e}^{\sum_j \lambda_j \Pi_j}$$

nekonzistentní vazby v důsledku šumu → MLME

- ML krok – najdu ML stav → $\{\boldsymbol{p}_j\}$
- ME krok – najdu $\hat{\rho}_{\mathrm{ME}}$ s pomocí konzistentních $\{\boldsymbol{p}_j\}$

alternativně řeším

$\max \; \mu \boldsymbol{E}(\rho) + \mathbf{log}\, \mathcal{L}(\rho)$ pro $\mu \to \mathbf{0}$

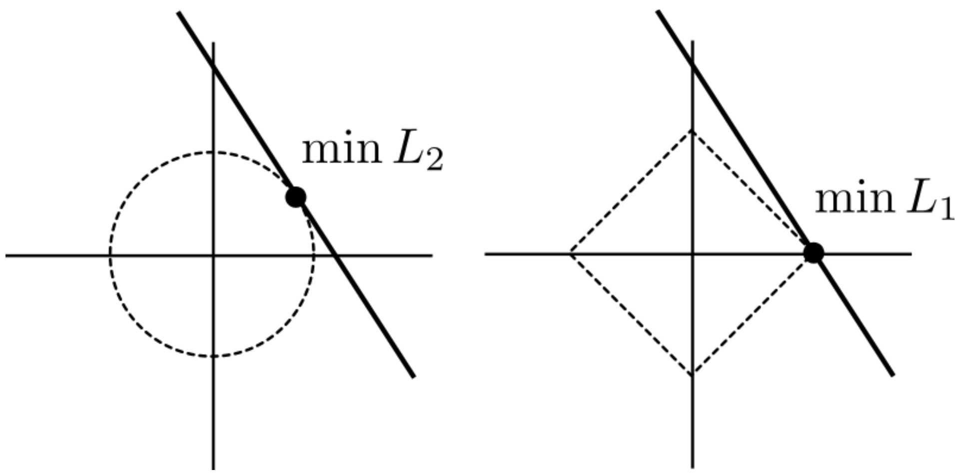
komprimované snímání

hledám “jednoduché/řídké” řešení

signál (vektor) je řídký v nějaké doméně

řídkost = malá L_0 norma

- není konvexní
- proto obvykle nahrazena L_1 normou s podobnými vlastnostmi



rekonstrukce kvantového stavu

$$\|\rho\|_1 = \sum_j s_j = \mathrm{Tr}(\rho) = 1$$ ... součet singulárních čísel

- norma nerozlišuje mezi kvantovými stavy
- rekonstrukce s pozitivitou se chová podobně jako komprimované snímání

optimalizace – konvexní programování

problém

$$\min f(x)$$
$$g_j(x) \leq 0, \quad j = 1, \ldots, r$$

globální minimum x^*

$$f^* = f(x^*)$$

Lagrangián

$$L(x, \mu) = f(x) + \sum_j \mu_j g_j(x)$$

(Lagrangeova) duální funkce

$$q(\mu) = \inf_x L(x, \mu)$$

- $q(\mu)$ je konkávní v μ i pro nekonvexní problém
- $q(\mu) \leq f^*, \quad \forall \mu > 0$

primární problém

$$\min f(x) \quad \text{s.t.} \quad g(x) \leq 0$$

duální problém

$$\max q(\mu) \quad s.t. \quad \mu \geq 0$$

<u>slabá dualita</u>

d^* ... optimum duálního problému

$d^* \leq f^*$ ... příklad max-min nerovnosti

$$d^* - f^*$$ ... duality gap

duální problém (snadný) poskytuje dolní mez pro původní problém

<u>silná dualita</u>

$$d^* = f^*$$

- neplatí obecně
- obvykle platí pro konvexní problémy

<u>Slaterova podmínka</u>

silná dualita platí pro konvexní problém

$$\min f(x),\ g_j(x) \le 0,\ j = 1, \ldots, r$$

pokud je problém striktně přípustný

$$\exists x:\ g_j(x) < 0,\ \ j = 1, \ldots, r$$

geometrická interpretace slabé duality:

množina $S = \{\ (\ g(x), f(x)\ )\ |\ x \in \mathbb{R}^n\ \}$

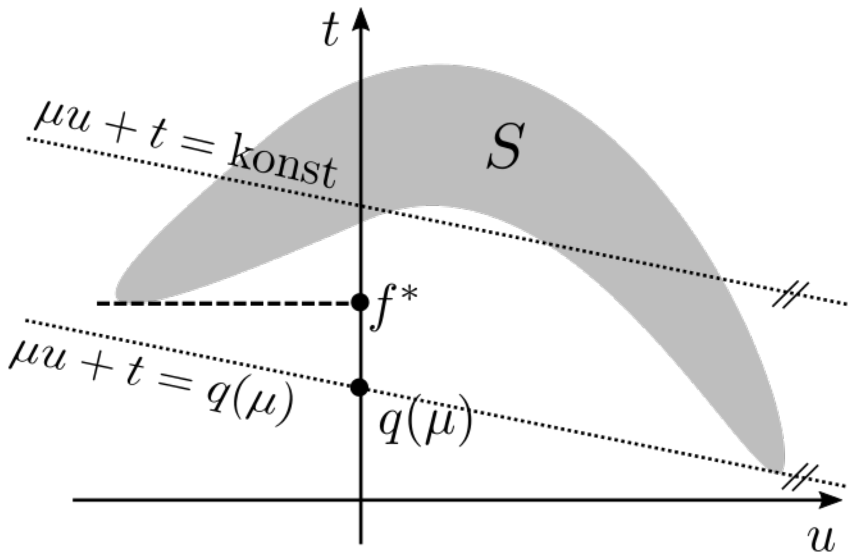
optimum $f^* = \inf\{\ t\ |\ (u, t) \in S,\ u \le 0$

duální funkce pro $\mu \ge 0$:

$$q(\mu) = \inf_{(u,t)\in S} \{\mu u + t\}$$

$\mu u + t = \text{konst}$ ... rovnice přímky $t(u) = \text{konst} - \mu u$

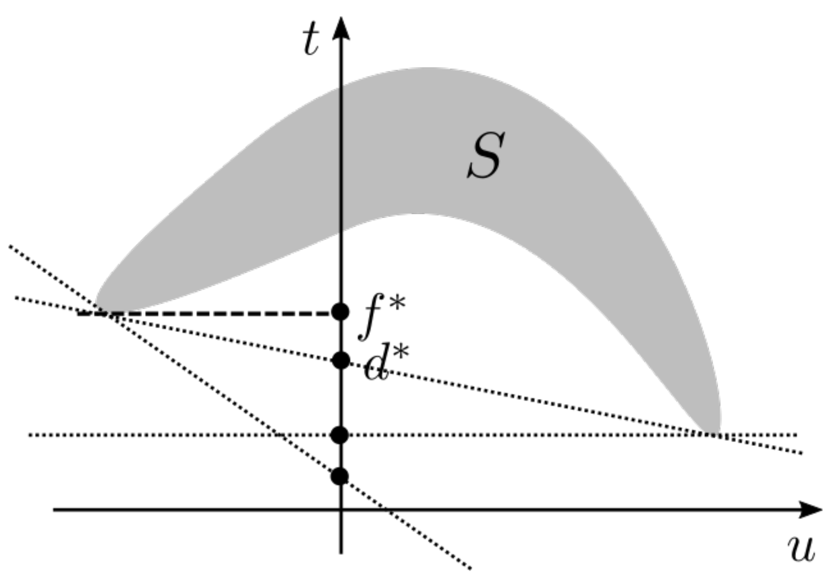
$u = 0 \ \Rightarrow\ t(0) = \mu u + t$ speciálně pro inf $t(0) = q(\mu)$



vidíme, že $q(\mu) \leq f^*$

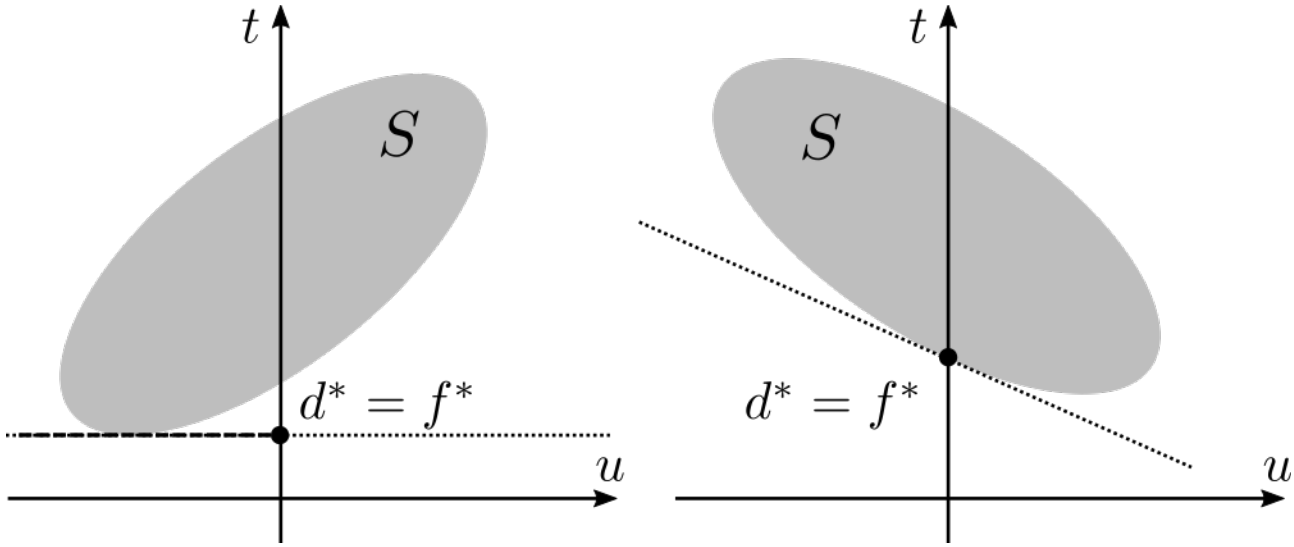
duální optimum d^*

$$d^* = \sup_{\mu \geq 0} q(\mu) = \sup_{\mu \geq 0} \inf_{(u,t) \in S} \{\mu u + t\}$$



vidíme, že $d^* \leq f^*$

geometrická interpretace silné duality:



komplementarita

použitím silné duality $f^* = d^*$ dostaneme

$$f(x^*) = q(\mu^*) = \inf_{x \in \mathbb{R}^n} \left[f(x) + \sum_j \mu_j^* g_j(x) \right] \leq L(x^*, \mu^*) \leq f(x^*)$$

complementary slackness

$$\mu_j^* g_j(x^*) = 0, \quad j = 1, \ldots, r$$

buď je vazba v optimu aktivní $g_j(x^*) = 0$ nebo je její Lagrangeův multiplikátor $\mu_j^* = 0$

splnění $L(x^*, \mu^*) = \inf_x L(x, \mu^*)$ vyžaduje (stacionarita)

$$\nabla_x L(x^*, \mu^*) = \nabla f(x^*) + \sum_j \mu_j^* \nabla g_j(x^*) = 0$$

Karush-Kuhn-Tucker (KKT) podmínky

konvexní f, g

pokud v x^* a μ^* platí KKT podmínky:

$$g_j(x^*) \leq 0 \quad \text{(přípustnost)}$$
$$\mu_j^* \geq 0$$

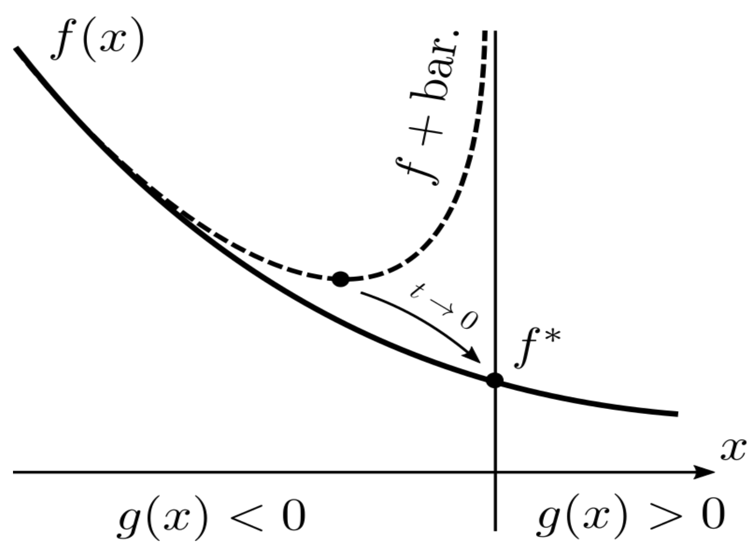
$$\mu_j^* g_j(x^*) = 0 \quad \text{(komplementarita)}$$

$$\nabla_x L(x^*, \mu^*) = \nabla f(x^*) + \sum_j \mu_j^* \nabla g_j(x^*) = 0 \quad \text{(stacionarita)}$$

pak x^* a μ^* jsou primární a duální optima a $d^* = f^*$

metoda vnitřních bodů (interior-point method)

$$\min f(x) - t \sum_j \log[-g_j(x)] \quad \text{(logaritmická bariéra)}$$



stacionarita

$$\nabla \boldsymbol{f}(\boldsymbol{x}^*) - \boldsymbol{t} \sum_j \frac{\mathbf{1}}{\boldsymbol{g}_j(\boldsymbol{x}^*)} \nabla \boldsymbol{g}_j(\boldsymbol{x}^*) = \mathbf{0} \qquad (1)$$

zavedeme

$$\mu_j^* = -\frac{\boldsymbol{t}}{\boldsymbol{g}_j(\boldsymbol{x}^*)} \geq \mathbf{0} \quad \Rightarrow \quad \mu_j^* \boldsymbol{g}_j(\boldsymbol{x}^*) = -\boldsymbol{t} \quad (2)$$

použitím perturbované komplementarity (2) v (1) dostaneme

$$\nabla \boldsymbol{f}(\boldsymbol{x}^*) + \sum_j \mu_j^* \nabla \boldsymbol{g}_j(\boldsymbol{x}^*) = \mathbf{0} \quad (3)$$

řešíme (2) a (3) kvazi-Newtonovou metodou

- vnitřní iterace

  pro dané $\boldsymbol{t}$ hledáme $\boldsymbol{x}^*(\boldsymbol{t})$

- vnější iterace

  snížování výšky bariéry

duality gap

podle (3) minimalizuje $\boldsymbol{x}^*$ Lagrangián $\boldsymbol{L}(\boldsymbol{x}, \mu^*)$:

$$\boldsymbol{q}(\mu^*) = \boldsymbol{L}(\boldsymbol{x}^*, \mu^*) = \boldsymbol{f}(\boldsymbol{x}^*) - \boldsymbol{r}\,\boldsymbol{t}$$

duality gap = $\boldsymbol{r}\,\boldsymbol{t}$