Optimalizace

Zárodek skript k předmětu A4B33OPT.
Text je neúplný a v průběhu semestru je doplňován a vylepšován.
Toto je verze ze dne **23. září 2011**.

Tomáš Werner

České vysoké učení technické
Fakulta elektrotechnická

Obsah

Kapitola 1

Úvod

1.1 Disciplína optimalizace

Optimalizace (přesněji matematická optimalizace) se zabývá minimalizací (či maximalizací) funkcí mnoha proměnných za případných omezujících podmínek. Do této definice se vejde překvapivé množství úloh z inženýrské praxe i přírodních věd: často přece chceme něco udělat 'nejlépe' v rámci 'daných možností'. Naučit se rozpoznávat optimalizační problémy kolem sebe je inženýrovi velmi užitečné. Optimalizace, též nazývaná *matematické programování*, je část aplikované matematiky, ležící na pomezí matematické analýzy, lineární algebry a informatiky. Je to moderní obor, který se rychle rozvíjí.

Příklady problémů, které vedou na optimalizační úlohy:

- Aproximuj naměřenou funkční závislost polynomem daného stupně, aby aproximace byla nejlepší.
- Investuj 1000 Kč do daných druhů akcií tak, aby očekávaný výnos byl velký a riziko malé.
- Rozmísti daný počet prodejen po městě tak, aby každý člověk měl do prodejny blízko.
- Najdi průběh řídícího signálu ruky robota tak, aby se dostala z místa A do místa B po dráze minimální délky (příp. minimálního času či výdaje energie) a bez kolize.
- Reguluj přívod plynu do kotle a oběh vody v topení tak, aby teplota v domě byla blízká kýžené teplotě.
- Navrhni plošný spoj daného zapojení, aby délka spojů (příp. počet průchodů mezi vrstvami) byla nejmenší.
- Najdi nejkratší cestu v počítačové síti.
- Vyhledej nejlepší spojení v jízdním řádu z místa A do místa B.
- Navrhni nejlepší školní rozvrh.
- Postav most o dané nosnosti při nejmenší spotřebě materiálu.
- Nauč umělou neuronovou síť.

Mimo inženýrskou praxi je optimalizace významná v přírodních vědách. Většinu fyzikálních zákonů lze formulovat tak, že nějaká veličina nabývá extrémální hodnoty. Živé organismy v každém okamžiku řeší, vědomě či podvědomě, množství optimalizačních úloh – např. se rozhodují pro nejlepší z možných chování.

V tomto kursu se nenaučíte řešit všechny tyto úlohy – už proto, že některé jsou velmi těžké. Ale naučíte se rozpoznat druh a obtížnost úloh a dostanete pevné základy pro přesné řešení těch snadnějších a přibližné řešení těch obtížnějších. Spektrum úloh, které dokážete řešit, se ještě podstatně rozšíří po absolvování navazujícího kursu *Kombinatorická optimalizace*.

Problém popsaný přirozeným jazykem je nejdříve nutno převést do jazyka matematiky – vytvořit jeho matematický *model* (tedy zvolit proměnné, účeloou funkci a omezení). Vytvoření matematického modelu vyžaduje zjednodušení reality, tj. zanedbání některých jevů (proto, že jsou buď 'nepodstatné' nebo že nejsou 'příliš podstatné' ale je obtížné je modelovat). Slova přirozeného jazyka jako 'nejlepší' nebo 'nejvýhodnější' mohou mít matematicky několik různých významů, tedy překlad slovní úlohy do matematického jazyka nemusí být jednoznačný. Konstrukce matematického modelu vyžaduje u složitějších úloh hodně práce a často několik cyklů pokus–omyl.

1.2 Značení

Standardní číselné množiny.

- $\mathbb{N}$ je množina přirozených čísel
- $\mathbb{Z}$ je množina celých čísel
- $\mathbb{Q}$ je množina racionálních čísel
- $\mathbb{R}$ je množina reálných čísel
- $\mathbb{R}_+$ resp. $\mathbb{R}_{++}$ je množina nezáporných resp. kladných reálných čísel.
- $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ je množina **rozšířených reálných čísel**, kde aritmetické operace jsou dodefinovány pro obě nekonečna (výrazy $\infty - \infty$, $0 \cdot \infty$, $0/0$ a ∞/∞ jsou nedefinovány).

Kartézský součin. Pro množiny $X_1, \ldots, X_n$ uspořádanou n-tici prvků $x_1 \in X_1, \ldots, x_n \in X_n$ označujeme kulatými závorkami $(x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n$. Kartézský součin n stejných množin značíme $X^n = X \times \cdots \times X$ (n-krát).

Množina definovaná vlastností prvků. Nechť $\varphi(x)$ je logický výrok s proměnnou $x \in X$, tedy $\varphi(x)$ je buď pravda nebo nepravda v závislosti na x. Symbol

$$\{\, x \in X \mid \varphi(x) \,\}$$

označuje množinu prvků množiny X, které splňují výrok $\varphi(x)$. Např. $\{\, x \in \mathbb{R} \mid -1 < x < 1 \,\}$ je otevřený interval $(-1, 1)$. Pokud z kontextu bude jasná množina X, budeme toto zkracovat na $\{\, x \mid \varphi(x) \,\}$, např. $\{\, x \mid -1 < x < 1 \,\}$.

Pokud $f\colon X \to Y$ je zobrazení z množiny X do množiny Y, budeme používat zkratku

$$\{\, y \in Y \mid y = f(x),\ x \in X,\ \varphi(x) \,\} = \{\, f(x) \mid x \in X,\ \varphi(x) \,\}$$

či pouze $\{\, f(x) \mid \varphi(x) \,\}$, pokud je X jasné z kontextu. Např. množina

$$\{\, y \in \mathbb{R} \mid y = x^2,\ x \in \mathbb{R},\ -1 < x < 1 \,\} = \{\, x^2 \mid x \in \mathbb{R},\ -1 < x < 1 \,\} = \{\, x^2 \mid -1 < x < 1 \,\}$$

je polouzavřený interval $\langle 0, 1)$.

Extrém reálné funkce na množině. Pro minimum *reálné* funkce f na množině X používáme označení

$$\min_{x\in X} f(x) = \min\{\, f(x) \mid x \in X \,\}.$$

Pro množinu prvků X, ve kterých se minimum nabývá, se používá symbol **argument minima**:

$$\operatorname*{argmin}_{x\in X} f(x) = \{\, x^* \in X \mid f(x^*) = \min_{x\in X} f(x) \,\}.$$

Uvědomte si, že $Y = \{\, f(x) \mid x \in X \,\} \subseteq \mathbb{R}$ je sama o sobě *množina*, je to obraz množiny X v zobrazení f. Tedy zatímco zápis $\min\{\, f(x) \mid x \in X \,\}$ je korektní a znamená $\min Y$, zápis $\operatorname{argmin}\{\, f(x) \mid x \in X \,\}$ je syntaktický nesmysl.

Příklad 1.1.

1. $\min_{x\in\mathbb{R}} |x-1| = \min\{\, |x-1| \mid x \in \mathbb{R} \,\} = 0$, $\operatorname*{argmin}_{x\in\mathbb{R}} |x-1| = \{1\}$
2. Nechť $(a_1, a_2, \ldots, a_5) = (1, 2, 3, 2, 3)$. Pak $\max_{i=1}^{5} a_i = 3$, $\operatorname*{argmax}_{i=1}^{5} a_i = \{3, 5\}$.
3. $\operatorname*{argmax}_{(x,y)\in\langle -5,5\rangle\times\mathbb{R}} x\cos y = \{\, (5, 2k\pi) \mid k \in \mathbb{Z} \,\} \cup \{\, (-5, (2k+1)\pi) \mid k \in \mathbb{Z} \,\}$ □

Cvičení

1.1. Najděte (úvahou) co nejjednodušší popis následujících množin:

a) $\{\, x^2 \mid x \in \mathbb{R} \,\}$
b) $\{\, 1/x \mid x \ge 1 \,\}$
c) $\{\, \mathrm{e}^{-x^2} \mid x \in \mathbb{R} \,\}$
d) $\{\, x+y \mid x^2+y^2 < 1 \,\}$
e) $\{\, x+y \mid x^2+y^2 = 1 \,\}$
f) $\{\, |x|+|y| \mid x^2+y^2 = 1 \,\}$
g) $\{\, x_1 + \cdots + x_n \mid \mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{x}\|_2 = 1 \,\}$
h) $\{\, |x-y| \mid x \in \langle 0,1\rangle,\ y \in (1,2\rangle \,\}$

1.2. Mějme množinu bodů v rovině $X = \langle -1, 1\rangle \times \{0\} = \{\, (x,0) \mid -1 \le x \le 1 \,\} \subseteq \mathbb{R}^2$. Načrtněte následující množiny:

a) $\left\{\, \mathbf{y} \in \mathbb{R}^2 \;\middle|\; \min_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|_2 \le 1 \,\right\}$
b) $\left\{\, \mathbf{y} \in \mathbb{R}^2 \;\middle|\; \max_{\mathbf{x}\in X} \|\mathbf{x}-\mathbf{y}\|_2 \le 2 \,\right\}$

1.3 Formulace optimalizačních úloh

Optimalizační úlohy lze formulovat jako hledání minima dané reálné funkce na dané množině:

$$\min_{x\in X} f(x). \tag{1.1}$$

V optimalizaci se často nazývá následující názvosloví. Funkce f se nazývá **účelová** (také pokutová, cenová, kriteriální) funkce. Množina X se nazývá množina **přípustných řešení**, což je vlastně protimluv, protože prvky X nejsou řešeními úlohy. Prvky množiny $\operatorname{argmin}_{x \in X} f(x)$ se pak nazývají **optimální řešení**.

Formulace (1.1) je velmi obecná a abstraktní, neboť množina X může být zcela libovolná. Existují tři široké kategorie úloh:

- Pokud je množina X konečná (i když typicky velmi velká), mluvíme o *kombinatorické optimalizaci*. Její prvky mohou být např. cesty v grafu, konfigurace Rubikovy kostky, nebo textové řetězce konečné délky. Příkladem je nalezení nejkratší cesty v grafu nebo problém obchodního cestujícího.
- Pokud množina X obsahuje reálná čísla či vektory, mluvíme o *spojité optimalizaci*. Příkladem je úloha lineárního programování.
- Pokud množina X obsahuje funkce, mluvíme o *variačním počtu*. Příkladem je nalézt tvar rovinné křivky, která při dané délce obepíná co největší obsah.

Tento kurs se věnuje téměř výhradně spojité optimalizaci. Zde prvky množiny X jsou n-tice reálných proměnných $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, tedy $X \subseteq \mathbb{R}^n$. Všimněte si, že zapisujeme $\mathbf{x}$ tučně, protože jej považujeme za vektor. Množina X je definována jako množina všech $\mathbf{x} \in \mathbb{R}^n$ splňujících rovnice

$$\begin{aligned} g_i(\mathbf{x}) &\leq 0, \quad i = 1, \ldots, m \\ h_i(\mathbf{x}) &= 0, \quad i = 1, \ldots, \ell \end{aligned} \tag{1.2}$$

pro dané reálné funkce n proměnných $g_1, \ldots, g_m$ a $h_1, \ldots, h_\ell$. V tomto případě se úloha zapisuje také jako

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{za podmínek} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, \ell. \end{aligned} \tag{1.3}$$

Příklad 1.2. Máme najít bod v rovině, který leží na kružnici s jednotkovým poloměrem a se středem v počátku a který je nejblíže danému bodu $\mathbf{a}$. Zde máme $n = 2$, $m = 0$, $\ell = 1$, $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$, $h_1(\mathbf{x}) = \|\mathbf{x}\| - 1$. To lze psát jako

$$\begin{aligned} \min \quad & \|\mathbf{a} - \mathbf{x}\| \\ \text{za podmínky} \quad & \|\mathbf{x}\| = 1 \end{aligned}$$

nebo jako $\min\{\, \|\mathbf{a} - \mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^2,\ \|\mathbf{x}\| = 1 \,\}$. □

Rovnice a nerovnice (1.2) jsou **omezující podmínky**, krátce **omezení**. Omezení tvaru $g_i(\mathbf{x}) \leq 0$ příp. $h_i(\mathbf{x}) = 0$ se nazývá omezení **typu nerovnosti** příp. **typu rovnosti**. Optimální řešení může být jedno, více, nebo nemusí existovat. Všimněte si, že omezení $h_i(\mathbf{x}) = 0$ je ekvivalentní dvěma omezením $h_i(\mathbf{x}) \leq 0$ a $-h_i(\mathbf{x}) \leq 0$, proto jsme vlastně omezení typu rovnosti v problému (1.3) nemuseli explicitně uvádět. Omezení mohou chybět ($m = \ell = 0$).

V závislosti na účelové funkci a omezeních může být nalezení optimálního řešení snadné, obtížné, nebo prakticky nemožné.

V tomto kursu se budeme zabývat těmito speciálními případy:

- $m = \ell = 0$, funkce f spojitá diferencovatelná.
  Vícerozměrná optimalizace bez omezení. Speciálně f kvadratická positivně definitní: lineární nejmenší čtverce.
- $m = 0$, $\ell > 0$, funkce f, h_i diferencovatelné.
  Zde se naučíte metodu Lagrangeových multiplikátorů.

- $m, \ell > 0$, funkce f, g_i, h_i direferencovatelné.
  Zobecnění metody Lagrangeových multiplikátorů: Karush-Kuhn-Tucker (KKT) podmínky.
- $m, \ell > 0$, funkce f, g_i, h_i afinní.
  Lineární programování (LP) a simplexová metoda.
- $m, \ell > 0$, funkce f kvadratická pozitivně semidefinitní, g_i, h_i afinní.
  Kvadratické programování (QP).
- $m, \ell > 0$, funkce f, g_i konvexní, h_i afinní.
  Konvexní programování, existuje jen jedno lokální minimum. Budeme se zabývat vlastnostmi konvexních množin a funkcí.

Cvičení

1.3. Formulujte (již však neřešte) následující úlohy ve tvaru $\min\{\, f(\mathbf{x}) \mid \mathbf{x} \in X \,\}$.

   a) Najdi dvojici bodů v rovině, z nichž jeden je uvnitř čtverce se středem v počátku a jednotkovou stranou a druhý je uvnitř kruhu se středem v bodě $(2, 2)$ a jednotkovým poloměrem, tak aby si tyto dva body byly nejblíže.
   b) Najdi dvě přirozená čísla se součtem 7 a nejmenším součinem.
   c) Nechť $\mathbf{A}$ je matice rozměru $m \times n$, kde $m < n$, a $\mathbf{b}$ je vektor délky m. Předpokládejte, že matice má hodnost m. Najdi řešení $\mathbf{x}$ soustavy $\mathbf{A}\mathbf{x} = \mathbf{b}$ tak, že délka vektoru $\mathbf{x}$ je minimální.
   d) Ve vesnici je n chalup o daných souřadnicích. Najdi souřadnice pošty tak, aby pošťák měl k nejvzdálenější (měřeno vzdušnou čarou) chalupě co nejblíže.
   e) Ve vesnici je n chalup o daných souřadnicích. Najdi souřadnice m pump tak, aby vzdálenost (vzdušnou čarou) od libovolné chalupy k nejbližší pumpě byla minimální.
   f) Je dáno m bodů v rovině. Najdi kružnici (s libovolným středem a poloměrem) tak, aby součet druhých mocnin vzdáleností bodů od kružnice byl minimální.

1.4. Vyřešte úvahou (příp. si nakreslete obrázek) následující úlohy:

   a) $\min\{\, x^2 + y^2 \mid x > 0,\ y > 0,\ xy \geq 1 \,\}$
   b) $\min\{\, (x-2)^2 + (y-1)^2 \mid x^2 \leq 1,\ y^2 \leq 1 \,\}$
   c) $\max\{\, \mathbf{x} \cdot \mathbf{y} \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{x} \cdot \mathbf{x} \leq 1 \,\}$ pro daný vektor $\mathbf{y} \in \mathbb{R}^n$ ($\mathbf{x} \cdot \mathbf{y}$ značí skalární součin). Zkuste nejprve pro $n = 1$, pak pro $n = 2$, pak zobecněte na libovolné n.
   d) Najdi rozměry krabice bez víka o jednotkovém objemu a co nejmenším povrchu.
   e) Hledá se n-tice čísel $x_1, \ldots, x_n \in \{-1, +1\}$ tak, že jejich součin je kladný a jejich součet minimální. Jako výsledek napište hodnotu tohoto minimálního součtu v závislosti na n.