OPT kompakt

sobota 9. května 2020 11:18

Optimálně dlouhá optimalizace. Aneb když chcete hlavní tvrzení a poznatky, bez zátěže důkazů.

Minilingebra

Rank je dimenze lin. obalu sloupců. Ale taky řádků:

$\operatorname{rank} \mathbf{A} = \operatorname{rank}(\mathbf{A}^T)$,

Plná hodnost: rank = min(m,n)
Rank(AB) <= min (rankA, rankB)

Některé vlastnosti determinantů:
- $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$
- $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$ (plyne z předchozího pro $\mathbf{B} = \mathbf{A}^{-1}$)

- $\det \mathbf{A}^T = \det \mathbf{A}$
- $\det \mathbf{A} = 0$ právě tehdy, když $\mathbf{A}$ je singulární

Soustava Ax = b:
Homogenní pokud b = 0, jinak nehomogenní

Věta 3.5. *Pro matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ jsou následující tvrzení ekvivalentní:*
1. $\operatorname{rng} \mathbf{A} = \mathbb{R}^m$
2. *Soustava $\mathbf{Ax} = \mathbf{y}$ má řešení pro každé $\mathbf{y}$.*
3. $\operatorname{rank} \mathbf{A} = m$
4. *Zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ je surjektivní, tj. $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^m$ (viz §1.1.2).*
5. *Řádky matice $\mathbf{A}$ jsou lineárně nezávislé.*
6. *Matice $\mathbf{A}$ má pravou inverzi, tj. existuje $\mathbf{B}$ tak, že $\mathbf{AB} = \mathbf{I}$.*
7. *Matice $\mathbf{AA}^T \in \mathbb{R}^{m\times m}$ je regulární.*

Věta 3.7. *Pro matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ jsou následující tvrzení ekvivalentní:*
1. $\operatorname{null} \mathbf{A} = \{\mathbf{0}\}$ *(tj. nulový prostor je triviální).*
2. *Soustava $\mathbf{Ax} = \mathbf{0}$ má jediné řešení $\mathbf{x} = \mathbf{0}$.*
3. $\operatorname{rank} \mathbf{A} = n$.
4. *Zobrazení $\mathbf{f}(\mathbf{x}) = \mathbf{Ax}$ je injektivní (viz §1.1.2).*
5. *Sloupce matice $\mathbf{A}$ jsou lineárně nezávislé.*
6. *Matice $\mathbf{A}$ má levou inverzi, tj. existuje $\mathbf{B}$ tak, že $\mathbf{BA} = \mathbf{I}$.*
7. *Matice $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{n\times n}$ je regulární.*

Pro každou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ platí

$$\dim \operatorname{rng} \mathbf{A} + \dim \operatorname{null} \mathbf{A} = n.$$

Věta 5.1. *Pro každou matici $\mathbf{A}$ platí*[2]

$$\operatorname{rng}(\mathbf{A}^T\mathbf{A}) = \operatorname{rng}(\mathbf{A}^T),$$
$$\operatorname{null}(\mathbf{A}^T\mathbf{A}) = \operatorname{null} \mathbf{A}.$$

Afinní kombinace vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ je lineární kombinace $\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k$, ve které koeficienty kombinace splňují

$$\alpha_1 + \cdots + \alpha_k = 1.$$

Afinní podprostor - množina uzavřená na afinní kombinace.

Vlastní čísla a vlastní vektory

Nechť pro čtvercovou matici $\mathbf{A} \in \mathbb{R}^{n\times n}$, nenulový vektor $\mathbf{v} \in \mathbb{C}^n$ a skalár $\lambda \in \mathbb{C}$ platí

$$\mathbf{Av} = \lambda\mathbf{v}.$$

Charakteristický polynom:

$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0,$

Ortogonalita

Skalární součin indukuje normu, společně dávají úhly:

$$\cos\varphi = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

Tedy vektory jsou ortogonální iff **x'y = 0**
Izometrie - zobrazení zachovávající úhly a vzdálenosti. **Ortogonální matice představuje izometrii.**
Zda je zobrazení izometrie můžeme ověřovat pomocí skalárního součinu: (Ax)'Ax = ... = x'x -> pak je izometrie.
Ortogonální doplňky:
Věta 4.1. *Pro každé podprostory $X, Y \subseteq \mathbb{R}^n$ platí:*
1. $\dim X + \dim X^\perp = n$,
2. $X \perp Y$ *a* $\dim X + \dim Y = n$ *implikuje* $Y = X^\perp$,

Nulový prostor je ortogonální doplněk (je to podprostor kolmý na podprostor lineárního obalu řádků matice):

$(\operatorname{rng}(\mathbf{A}^T))^\perp = \operatorname{null} \mathbf{A}.$

QR rozklad:
Věta 4.6. *Každou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ lze rozložit na součin*

$$\mathbf{A} = \mathbf{QR},$$

kde matice $\mathbf{Q} \in \mathbb{R}^{m\times m}$ je ortogonální a matice $\mathbf{R} \in \mathbb{R}^{m\times n}$ je horní trojúhelníková.
Redukovaný QR rozklad:
Vynecháme-li tedy z matice $\mathbf{R}$ posledních $m-n$ řádků a z matice $\mathbf{Q}$ posledních $m-n$ sloupců, součin $\mathbf{QR}$ se nezmění. Pro $m > n$ lze tedy každou $\mathbf{A} \in \mathbb{R}^{m\times n}$ rozložit jako $\mathbf{A} = \mathbf{QR}$ kde $\mathbf{Q} \in \mathbb{R}^{m\times n}$ má ortonormální sloupce a $\mathbf{R} \in \mathbb{R}^{n\times n}$ je horní trojúhelníková.
Př použití: řešení soustav rovnic (využívá se regularity Q)

Ortogonální projekce:
Matice U obsahuje ve sloupcích ortonormální bázi podprostoru:

$\mathbf{x} = \mathbf{UU}^T\mathbf{z}$

$\mathbf{P} = \mathbf{UU}^T$ => $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^T$

Projekce na ortogonální doplněk:
x = (I - UUT)z
Vzdálenost bodu x od **ortogonálního doplňku** podprostoru s bází U je:

$\|\mathbf{U}^T\mathbf{z}\|$

Pokud máme podprostor s obecnou bází A:

$\mathbf{P} = \mathbf{AA}^+ = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T.$

Spektrální rozklad

V .. matice vektorů příslušných vlastním číslům, Λ .. matice s vlastními čísly na diagonále.
Pokud je $\mathbf{V}$ regulární (tj. matice $\mathbf{A}$ má n lineárně nezávislých vlastních vektorů), je invertovatelná a (6.5) lze psát jako

$$\mathbf{A} = \mathbf{V\Lambda V}^{-1}. \tag{6.6}$$

Vztahu (6.6) se pak říká **rozklad matice podle vlastních čísel** nebo **spektrální rozklad**. V tom případě je matice $\mathbf{A}$ podobná[4] diagonální matici (neboli **diagonalizovatelná**), protože z (6.6) plyne $\mathbf{V}^{-1}\mathbf{AV} = \mathbf{\Lambda}$.

Věta 6.1. *Nechť $\mathbf{A} \in \mathbb{R}^{n\times n}$. Následující dvě tvrzení jsou ekvivalentní:*
- *Matice $\mathbf{A}$ je symetrická.*
- *Všechna vlastní čísla matice $\mathbf{A}$ jsou reálná a $\mathbf{A}$ má n vlastních vektorů které jsou po dvojicích ortogonální.*

Tedy spektrální rozklad **symetrické matice: A = VΛVT**

Vztah hodnosti matice A a Λ:

$\operatorname{rank} \mathbf{\Lambda} = \operatorname{rank} \mathbf{A}.$

Nejmenší čtverce (čtverečky)

Obecná úloha řešení **přeurčené** (tzn. bez řešení) soustavy rovnic přibližně kritériem nejmenších čtverců:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|^2.$$

Minimální velikost: vektor **Ax - b** musí být komý na každý sloupec A (intuitivně), tedy:

$\mathbf{A}^T(\mathbf{Ax} - \mathbf{b}) = \mathbf{0},$

Tato soustava (**normální rovnice**) má vždy řešení (díky větě 5.1 o kousek výše v minilingebře).

Pokud je matice ATA **regulární (A má lin. nez. sloupce)**, řešení:

$\mathbf{x} = \mathbf{A}^+\mathbf{b},$

A+ je **pseudoinverze** (a jedna z levých inverzí A):

$\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T.$

Pokud A má **lin. záv. sloupce**, pak má úloha **nekonečně mnoho řešení.**

Častěji řešíme pomocí **QR rozkladu (A = QR).** Po dosazení do normální rovnice a úpravách:

$\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}.$

Matice A musí být **čtvercová a s lin. nez. sloupci.**

Lineární regrese
Minimum výrazu:

$\|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2,$

"Minimalizujeme rozdíl pozorovaných měření y a predikcí modelu pro danou hodnotu x."

Řešení nedourčené soustavy (tzn. má nekonečně mnoho řešení)
Kritérium minimalizace normy řešení:

Kvadratické formy, funkce a jejich extrémy

Homogenní polynom = všechny jeho monomy mají stejný stupeň (monomy jsou to, co získáme, pokud rozsekáme polynom podle sčítání/odčítání)
Kvadratická forma na $\mathbb{R}^n$ je homogenní polynom $f\colon \mathbb{R}^n \to \mathbb{R}$ druhého stupně. Je pohodlné ji zapsat v maticovém tvaru
Pozn. je zvykem předpokládat, že je A symetrická. (Kdyby nebyla, můžeme ji nahradit její antisymetrickou částí):

$$\mathbf{A} = \underbrace{\tfrac{1}{2}(\mathbf{A} + \mathbf{A}^T)}_{\text{symetrická}} + \underbrace{\tfrac{1}{2}(\mathbf{A} - \mathbf{A}^T)}_{\text{antisymetrická}}.$$

Definitnost kvadratické normy a její matice:
Čtvercovou matici $\mathbf{A}$ nazýváme
- **positivně** [**negativně**] **semidefinitní**, když pro každé $\mathbf{x}$ platí $\mathbf{x}^T\mathbf{Ax} \geq 0$ [$\mathbf{x}^T\mathbf{Ax} \leq 0$],
- **positivně** [**negativně**] **definitní**, když pro každé $\mathbf{x} \neq \mathbf{0}$ platí $\mathbf{x}^T\mathbf{Ax} > 0$ [$\mathbf{x}^T\mathbf{Ax} < 0$],
- **indefinitní**, když existuje $\mathbf{x}$ a $\mathbf{y}$ tak, že $\mathbf{x}^T\mathbf{Ax} > 0$ a $\mathbf{y}^T\mathbf{Ay} < 0$.

Matice může mít i několik těchto vlastností najednou. Např. positivně definitní matice je zároveň positivně semidefinitní. Nulová matice je zároveň positivně i negativně semidefinitní.
Vztah definitnosti k extrémům:
Věta 6.2. *Nechť funkce f je dána jako $f(\mathbf{x}) = \mathbf{x}^T\mathbf{Ax}$.*
- *Je-li $\mathbf{A}$ positivně [negativně] semidefinitní, pak f v bodě $\mathbf{0}$ nabývá minimum [maximum].*
- *Je-li $\mathbf{A}$ positivně [negativně] definitní, pak f v bodě $\mathbf{0}$ nabývá ostré minimum [maximum].*
- *Je-li $\mathbf{A}$ indefinitní, pak f nemá minimum ani maximum.*

Kritéria definitnosti (VŠECHNY PRO SYMETRICKOU MATICI - MUSÍME SYMETRIZOVAT!):
- **Hlavní minor** matice $\mathbf{A}$ je číslo $\det \mathbf{A}_I$ pro nějakou neprázdnou $I \subseteq \{1, \dots, n\}$.

$\|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2$,

"Minimalizujeme rozdíl pozorovaných měření y a predikcí modelu pro danou hodnotu x."

Řešení nedourčené soustavy (tzn. má nekonečně mnoho řešení)

Kritérium minimalizace normy řešení:

$$\min\{\, \|\mathbf{x}\|^2 \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} = \mathbf{b}\,\}.$$

Řešení: vektor kolmý na podprostor nullA.

$$\mathbf{x} = \mathbf{A}^+\mathbf{b},$$

$$\mathbf{A}^+ = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$$

A+ je **pseudoinverze** matice A s **lin. nez. řádky.** Je to jedna z pravých inverzí A.

Použití spektrálního rozkladu a SVD rozklad

Stopa čtvercové matice - součet diagonálních prvků

vlastnosti:

- $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr}\mathbf{A} + \operatorname{tr}\mathbf{B}$
- $\operatorname{tr}(\alpha\mathbf{A}) = \alpha \operatorname{tr}\mathbf{A}$
- $\operatorname{tr}(\mathbf{A}^T) = \operatorname{tr}\mathbf{A}$
- $\operatorname{tr}(\mathbf{A}\mathbf{B}) = \operatorname{tr}(\mathbf{B}\mathbf{A})$ pro každé $\mathbf{A} \in \mathbb{R}^{m\times n}$ a $\mathbf{B} \in \mathbb{R}^{n\times m}$ (tzv. *cykličnost stopy*)

$\operatorname{tr}\mathbf{A} = \lambda_1 + \cdots + \lambda_n$,

Frobeniova norma:

$$\|\mathbf{A}\| = \sqrt{\langle \mathbf{A}, \mathbf{A}\rangle} = \sqrt{\operatorname{tr}(\mathbf{A}^T\mathbf{A})} = \sqrt{\operatorname{tr}(\mathbf{A}\mathbf{A}^T)}$$

V dalších bodech je předpokládáno vzestupné seřazení vlastních čísel.

Úloha na minimální stopu:

Věta 7.3. *Nechť $k \le n$. Platí*[4]

$$\min\{\, \operatorname{tr}(\mathbf{X}^T\mathbf{A}\mathbf{X}) \mid \mathbf{X} \in \mathbb{R}^{n\times k},\ \mathbf{X}^T\mathbf{X} = \mathbf{I}\,\} = \lambda_1 + \cdots + \lambda_k$$

a minimum se nabývá pro $\mathbf{X} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k]$.

Proložení bodů podprostorem:

$$\min\{\, \|\mathbf{A}\mathbf{X}\|^2 \mid \mathbf{X} \in \mathbb{R}^{n\times(n-k)},\ \mathbf{X}^T\mathbf{X} = \mathbf{I}\,\}.$$

Protože $\|\mathbf{A}\mathbf{X}\|^2 = \operatorname{tr}(\mathbf{X}^T\mathbf{A}^T\mathbf{A}\mathbf{X})$, úlohu (7.20) můžeme vyřešit pomocí Věty 7.3. Spočítáme spektrální rozklad $\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T = \mathbf{A}^T\mathbf{A}$ matice $\mathbf{A}^T\mathbf{A}$ (kde opět předpokládáme, že diagonální prvky matice $\boldsymbol{\Lambda}$ jsou vzestupně seřazeny). Rozdělíme-li (ortogonální) matici $\mathbf{V} = [\mathbf{X} \ \ \mathbf{Y}] \in \mathbb{R}^{n\times n}$ do dvou bloků $\mathbf{X} \in \mathbb{R}^{n\times(n-k)}$ a $\mathbf{Y} \in \mathbb{R}^{n\times k}$, pak $\mathbf{X}$ je řešení úlohy (7.20) a (viz §4.6.1) sloupce $\mathbf{Y}$ tvoří ortonormální bázi hledaného podprostoru (7.19). Tedy:

- Posledních k sloupců matice $\mathbf{V}$ (tj. ortonormální vlastní vektory matice $\mathbf{A}^T\mathbf{A}$ odpovídající k největším vlastním číslům) tvoří ortonormální bázi lineárního podprostoru dimenze k, který minimalizuje součet čtverců vdáleností k bodům $\mathbf{a}_1, \ldots, \mathbf{a}_m$.
- Prvních $n-k$ sloupců matice $\mathbf{V}$ (tj. ortonormální vlastní vektory matice $\mathbf{A}^T\mathbf{A}$ odpovídající $n-k$ nejmenším vlastním číslům) tvoří ortonormální bázi ortogonálního doplňku tohoto podprostoru.

Nejbližší matice nižší hodnosti (ekvivalentní úloze výše):

$$\min\{\, \|\mathbf{A} - \mathbf{B}\|^2 \mid \mathbf{B} \in \mathbb{R}^{m\times n},\ \operatorname{rank}\mathbf{B} \le k\,\}.$$

Vyřešíme předchozí úlohu a pak:

$$\mathbf{B} = \mathbf{A}\mathbf{Y}\mathbf{Y}^T = \mathbf{A}(\mathbf{I} - \mathbf{X}\mathbf{X}^T),$$

Nebo:

Věta 7.6 (Eckart-Young). *Nechť $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ je SVD matice $\mathbf{A}$. Řešení úlohy (7.22) je*

$$\mathbf{B} = \mathbf{U}\mathbf{S}'\mathbf{V}^T = s_{p-k+1}\mathbf{u}_{p-k+1}\mathbf{v}_{p-k+1}^T + \cdots + s_p\mathbf{u}_p\mathbf{v}_p^T, \tag{7.32}$$

kde matice $\mathbf{S}'$ se získá z matice $\mathbf{S}$ vynulováním $p-k$ nejmenších diagonálních prvků (v sumě dyád tedy předpokládáme[2] *$0 \le s_1 \le \cdots \le s_p$).*

Optimální hodnota obou úloh:

$\lambda_1 + \cdots + \lambda_{n-k}$,

Proložení afinním podprostorem:

Odečtením těžiště od bodů převedeme na úlohu výše, vyřešíme a pak zase přičteme těžiště.

Přeurčená homogenní lineární soustava

Přeurčená zde znamená, že má **pouze triviální řešení.** To nechceme, tak přidáme požadavek:

$$\min\{\, \|\mathbf{A}\mathbf{x}\| \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{x}^T\mathbf{x} = 1\,\}.$$

Řešení: vlastní vektor matice ATA příslušný jejímu nejmenšímu vlastnímu číslu.

Singulární rozklad

Věta 7.5. *Každou matici $\mathbf{A} \in \mathbb{R}^{m\times n}$ lze rozložit jako*

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T = s_1\mathbf{u}_1\mathbf{v}_1^T + \cdots + s_p\mathbf{u}_p\mathbf{v}_p^T. \tag{7.30}$$

kde matice $\mathbf{S} \in \mathbb{R}^{m\times n}$ je diagonální s diagonálními prvky $s_1, \ldots, s_p$ (kde $p = \min\{m, n\}$) a matice $\mathbf{U} = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m] \in \mathbb{R}^{m\times m}$ a $\mathbf{V} \in \mathbb{R}^{n\times n} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$ jsou ortogonální.

rankA = rankS

Redukovaný rozklad:

$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, kde $\mathbf{S} = \operatorname{diag}(s_1, \ldots, s_p) \in \mathbb{R}^{p\times p}$ je diagonální a $\mathbf{U} = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_p] \in \mathbb{R}^{m\times p}$ a $\mathbf{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_p] \in \mathbb{R}^{n\times p}$ mají ortonormální sloupce.

Nenulová singulární čísla matice A jsou druhé odmocniny nenulových vkastních čísel matic ATA a AAT.

Lokální extrémy vázané rovnostmi

Nejmenší čtverce s omezeními:

$$\begin{aligned} \min \quad & \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 \\ \text{za podmínek} \quad & \mathbf{C}\mathbf{x} = \mathbf{d} \end{aligned}$$

Řešení:

$$\begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T\mathbf{b} \\ \mathbf{d} \end{bmatrix}.$$

Tvrzení 10.2. *Nechť je zobrazení $\mathbf{g}$ spojitě diferencovatelné v bodě $\mathbf{x} \in X$. Nechť vektor $\mathbf{v} \in \mathbb{R}^n$ je tečný k množině X v bodě $\mathbf{x}$. Pak $\mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$.*

- *Je-li $\mathbf{A}$ positivně [negativně] definitní, pak f v bodě $\mathbf{0}$ nabývá ostré minimum [maximum].*
- *Je-li $\mathbf{A}$ indefinitní, pak f nemá minimum ani maximum.*

Kritéria definitnosti (VŠECHNY PRO SYMETRICKOU MATICI - MUSÍME SYMETRIZOVAT!):

- **Hlavní minor** matice $\mathbf{A}$ je číslo $\det \mathbf{A}_I$ pro nějakou neprázdnou $I \subseteq \{1, \ldots, n\}$.
- **Vůdčí hlavní minor** matice $\mathbf{A}$ je její hlavní minor pro nějaké $I = \{1, \ldots, k\}$ a $1 \le k \le n$.

Věta 6.3. *Symetrická matice je*

- *positivně definitní, právě když všechny její vůdčí hlavní minory jsou kladné,*
- *positivně semidefinitní, právě když všechny její hlavní minory jsou nezáporné.*

Věta 6.4. *Symetrická matice je*

- *positivně [negativně] semidefinitní, právě když má všechna vlastní čísla nezáporná [nekladná]*
- *positivně [negativně] definitní, právě když má všechna vlastní čísla kladná [záporná]*
- *indefinitní, právě když má alespoň jedno kladné a alespoň jedno záporné vlastní číslo.*

Kvadratická funkce:

Kvadratická funkce je polynom (ne nutně homogenní) druhého stupně. Lze jej psát v maticovém tvaru[5]

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c, \tag{6.12}$$

Doplnění na čtverec:

Někdy lze najít $\mathbf{x}_0 \in \mathbb{R}^n$ a $y_0 \in \mathbb{R}$ takové, že

$$\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = (\mathbf{x} - \mathbf{x}_0)^T\mathbf{A}(\mathbf{x} - \mathbf{x}_0) + y_0.$$

$$\mathbf{b} = -2\mathbf{A}\mathbf{x}_0,$$
$$c = \mathbf{x}_0^T\mathbf{A}\mathbf{x}_0 + y_0,$$

Když má první rovnice řešení (můžeme invertovat matici), pak lye najít doplnění na čtverec. Pokud nemá, doplnění není možné.

Extrémy:

Pokud doplnění na čtverec **existuje**, extrém se vyšetří **jako u kvadratické formy**.

Pokud **neexistuje**, pak kvadratická funkce extrém **nemá**.

Vrstevnice kvadratické funkce se nazývá **kvadrika**. Tedy kvadrika je množina[7]

$$\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = 0\,\} \tag{6.17}$$

všech řešení kvadratické rovnice, neboli množina všech kořenů kvadratické funkce.

Mikroanalýza

Tabulka derivací matic:

zobrazení	(totální) derivace	poznámka
$\mathbf{f}(\mathbf{x}) = \mathbf{x}$	$\mathbf{f}'(\mathbf{x}) = \mathbf{I}$	$\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^n$
$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$	$\mathbf{f}'(\mathbf{x}) = \mathbf{A}$	$\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m,\ \mathbf{A} \in \mathbb{R}^{m\times n},\ \mathbf{b} \in \mathbb{R}^m$
$f(\mathbf{x}) = \mathbf{a}^T\mathbf{x}$	$f'(\mathbf{x}) = \mathbf{a}^T$	$f\colon \mathbb{R}^n \to \mathbb{R},\ \mathbf{a} \in \mathbb{R}^n$
$f(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$	$f'(\mathbf{x}) = 2\mathbf{x}^T$	$f\colon \mathbb{R}^n \to \mathbb{R}$
$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$	$f'(\mathbf{x}) = \mathbf{x}^T(\mathbf{A} + \mathbf{A}^T)$	$f\colon \mathbb{R}^n \to \mathbb{R},\ \mathbf{A} \in \mathbb{R}^{n\times n}$
$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{g}(\mathbf{x})$	$\mathbf{f}'(\mathbf{x}) = \mathbf{A}\mathbf{g}'(\mathbf{x})$	$\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m,\ \mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^l,\ \mathbf{A} \in \mathbb{R}^{m\times l}$
$f(\mathbf{x}) = \|\mathbf{x}\|$	$f'(\mathbf{x}) = \mathbf{x}^T/\|\mathbf{x}\|$	$f\colon \mathbb{R}^n \to \mathbb{R}$
$f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\mathbf{g}(\mathbf{x})$	$f'(\mathbf{x}) = 2\mathbf{g}(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$	$f\colon \mathbb{R}^n \to \mathbb{R},\ \mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$
$f(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\mathbf{h}(\mathbf{x})$	$f'(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T\mathbf{h}'(\mathbf{x}) + \mathbf{h}(\mathbf{x})^T\mathbf{g}'(\mathbf{x})$	$f\colon \mathbb{R}^n \to \mathbb{R},\ \mathbf{g}, \mathbf{h}\colon \mathbb{R}^n \to \mathbb{R}^m$

Směrová derivace:

Věta 8.4. *Nechť zobrazení $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ je diferencovatelné v bodě $\mathbf{x}$. Pak jeho směrová derivace v bodě $\mathbf{x}$ ve směru $\mathbf{v}$ je rovna $\mathbf{f}'(\mathbf{x})\mathbf{v}$.*

Pozn. v OPT se povoluje nenormalizovaný směr **v**.

Parciální derivace jsou vlastně směrové derivace se směrem udávaným vektory kanonické báze

Vlastnosti gradientu:

- Směrová derivace je největší ve směru $\mathbf{v} = \nabla f(\mathbf{x})/\|\nabla f(\mathbf{x})\|$, tedy když je $\mathbf{v}$ rovnoběžný s gradientem a stejně orientovaný. Tedy směr gradientu je *směr největšího růstu* funkce.
- Velikost gradientu $\|\nabla f(\mathbf{x})\|$ je strmost funkce ve směru největšího růstu.
- Směrová derivace ve směru kolmém na gradient je nulová.

Taylorův polynom druhého stupně:

$$T_2(\mathbf{y}) = f(\mathbf{x}) + f'(\mathbf{x})\,(\mathbf{y} - \mathbf{x}) + \tfrac{1}{2}(\mathbf{y} - \mathbf{x})^T f''(\mathbf{x})\,(\mathbf{y} - \mathbf{x}).$$

Hessova matice a extrémy:

Věta 9.5. *Nechť $f\colon \mathbb{R}^n \to \mathbb{R}$ a $X \subseteq \mathbb{R}^n$. Nechť $\mathbf{x}$ je vnitřní bod množiny X. Nechť je funkce f v bodě $\mathbf{x}$ dvakrát diferencovatelná a platí $f'(\mathbf{x}) = \mathbf{0}$. Pak:*

- *Je-li $\mathbf{x}$ lokální minimum [maximum] funkce f na množině X, pak Hessova matice $f''(\mathbf{x})$ je positivně [negativně] semidefinitní.*
- *Je-li $f''(\mathbf{x})$ positivně [negativně] definitní, pak $\mathbf{x}$ je ostré lokální minimum [maximum] funkce f na množině X.*
- *Je-li $f''(\mathbf{x})$ indefinitní, pak $\mathbf{x}$ není lokální extrém funkce f na množině X.*

POZOR na směry implikací! Pokud je matice semidefinitní, o extrému nemůžeme nic říct!

Iterační metody na hledání volného lokálního minima funkce

Obecný tvar:

Iterační metody na hledání lokálního minima spojité funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ mají tvar

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{v}_k, \tag{9.1}$$

kde vektor $\mathbf{v}_k \in \mathbb{R}^n$ je **směr hledání** a skalár $\alpha_k > 0$ je **délka kroku**. V **sestupných metodách** (*descent methods*) hodnota účelové funkce monotonně klesá[2], $f(\mathbf{x}_{k+1}) < f(\mathbf{x}_k)$.

Podmínka sestupnosti směru:

Název metody	co počítá	iterační vzorec	poznámka
Gradientní	$\min f(\mathbf{x})$ pro $f\colon \mathbb{R}^n \to \mathbb{R}$	$\mathbf{x} = \mathbf{a} - f'(\mathbf{a})^T$	směr největšího spádu
Newtonova	$\mathbf{g}(\mathbf{x}) = \mathbf{o}$ pro $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^n$	$\mathbf{x} = \mathbf{a} - (\mathbf{g}'(\mathbf{a}))^{-1}\,\mathbf{g}(\mathbf{a})$	$\mathbf{g}'(a) \in \mathbb{R}^{n\times n}$ regul.
Newtonova	$\min f(\mathbf{x})$ pro $f\colon \mathbb{R}^n \to \mathbb{R}$	$\mathbf{x} = \mathbf{a} - (f''(\mathbf{a}))^{-1}\,(f'(\mathbf{a}))^T$	$(f')^T = \mathbf{g}$
Gauss-Newtonova	$\min \|\mathbf{g}(\mathbf{x})\|^2$ pro $g\colon \mathbb{R}^n \to \mathbb{R}^m$	$\mathbf{x} = \mathbf{a} - (\mathbf{g}'(\mathbf{a}))^{+}\,\mathbf{g}(\mathbf{a})$	$(\mathbf{g}')^+ = ((\mathbf{g}')^T\mathbf{g}')^{-1}(\mathbf{g}')^T$
Levenberg-Marq.	$\min \|\mathbf{g}(\mathbf{x})\|^2$ pro $g\colon \mathbb{R}^n \to \mathbb{R}^m$	$\mathbf{x} = \mathbf{a} - (\mathbf{g}'(\mathbf{a}))^{*}\,\mathbf{g}(\mathbf{a})$	$(\mathbf{g}')^* = ((\mathbf{g}')^T\mathbf{g}' + \mu_k\mathbf{I})^{-1}(\mathbf{g}')^T$

Podmínka sestupnosti Newtonovy metody na minimalizaci funkce:

Toto platí, když $f'(\mathbf{x}_k) \neq \mathbf{0}$ (tj. $\mathbf{x}_k$ není stacionární bod) a matice $f''(\mathbf{x}_k)$ je positivně definitní

Konvexní množiny, funkce

Konvexní množiny:

$$\begin{bmatrix} \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{d} \end{bmatrix}.$$

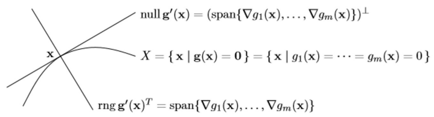
Tvrzení 10.2. *Nechť je zobrazení $\mathbf{g}$ spojitě diferencovatelné v bodě $\mathbf{x} \in X$. Nechť vektor $\mathbf{v} \in \mathbb{R}^n$ je tečný k množině X v bodě $\mathbf{x}$. Pak $\mathbf{g}'(\mathbf{x})\mathbf{v} = \mathbf{0}$.*

Podmínky pro extrémy:

Tvrzení 10.4. *Nechť $\mathbf{x} \in X$ je lokální extrém funkce f na množině X. Nechť f a $\mathbf{g}$ jsou v bodě $\mathbf{x}$ spojitě diferencovatelné. Nechť $\mathbf{v} \in \mathbb{R}^n$ je tečný vektor k množině X v bodě $\mathbf{x}$. Pak $f'(\mathbf{x})\mathbf{v} = 0$.*

Věta 10.5. *Nechť $\mathbf{x} \in X$ je lokální extrém funkce f na množině X. Nechť f a $\mathbf{g}$ jsou v bodě $\mathbf{x}$ spojitě diferencovatelné. Nechť platí (10.12). Pak gradient $\nabla f(\mathbf{x}) = f'(\mathbf{x})^T$ patří do ortogonálního prostoru množiny X v bodě $\mathbf{x}$, tj.*

$$\nabla f(\mathbf{x}) \in \operatorname{span}\{\nabla g_1(\mathbf{x}), \dots, \nabla g_m(\mathbf{x})\}. \qquad (10.14)$$



Lagrangeovy multiplikátory

Zaveďme **Lagrangeovu funkci** $L\colon \mathbb{R}^{n+m} \to \mathbb{R}$ jako

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \cdots + \lambda_m g_m(\mathbf{x}).$$

Funkci zderivujeme a položíme rovnu nule.
Poz. **Hessova** matice Lagrangeovy funkc je **vždy indefinitní,** neboli celkem k ničemu

Lineární programování

Minimalizace (maximalizace) lineární funkce za omezujících podmínek ve tvaru lineárních rovnic a nerovnic.

Speciální tvary:
Maticový:

$$\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} \geq \mathbf{b} \,\},$$

Rovnicový:

$$\min\{\, \mathbf{c}^T\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n,\ \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,\}.$$

Rovnicový získáme:

- Nerovnost $\mathbf{a}^T\mathbf{x} \geq b$ nahradíme dvěma omezeními $\mathbf{a}^T\mathbf{x} - u = b$, $u \geq 0$. Pomocné proměnné u se říká **slacková proměnná**[3]. Podobně bychom převedli nerovnost $\mathbf{a}^T\mathbf{x} \leq b$ na rovnost.
- Neomezenou proměnnou $x \in \mathbb{R}$ rozdělíme na dvě nezáporné proměnné $x^+ \geq 0$, $x^- \geq 0$ přidáním podmínky $x = x^+ - x^-$.

Užitečné vztahy:

$$\max_i a_i \leq b \iff (\forall i)(a_i \leq b).$$

min max| .. | -> min z : -z <= .. <= z

Normy
Funkce ||.|| z R^n -> R se nazývá **vektorová norma** iff:

1. Jestliže $\|\mathbf{x}\| = 0$ pak $\mathbf{x} = \mathbf{0}$.
2. $\|\alpha\mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|$ pro každé $\alpha \in \mathbb{R}$ a $\mathbf{x} \in \mathbb{R}^n$ (norma je kladně homogenní).
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ pro každé $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (trojúhelníková nerovnost).

Manhattanská norma:

$\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_n|.$

Eukleidovská norma:

$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\mathbf{x}^T\mathbf{x}}.$

Čebyševova max-norma:

$\|\mathbf{x}\|_\infty = \lim_{p\to\infty} \|\mathbf{x}\|_p = \max\{|x_1|, \dots, |x_n|\}$

Konvexní množiny, funkce

Konvexní množina:

Množina $X \subseteq \mathbb{R}^n$ je konvexní, pokud pro všechna $\mathbf{x}, \mathbf{y} \in X$ a libovolné $\alpha \in [0,1]$ platí $(1-\alpha)\mathbf{x} + \alpha\mathbf{y} \in X$.

Průnik konvexních množin je konvexní množina.

Konvexní funkce:

Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$ jestliže pro všechna $\mathbf{x}, \mathbf{y} \in X$ a každé $\alpha \in [0,1]$ platí

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y})$$

Konvexní optimalizační úloha: optimalizujeme **konvexní** účelovou funkci na **konvexní množině.**

Pro funkci $f\colon \mathbb{R}^n \to \mathbb{R}$ definujeme:

- **Epigraf** funkce je množina $\{\, (\mathbf{x}, y) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}) \leq y \,\}$.
- **Subkontura**[1] výšky y je množina $\{\, \mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq y \,\}$.

Funkce je konvexní iff její epigraf je konvexní množina.
Každá subkontura konvexní funkce je konvexní množina. (Pozor na implikaci)

Věta 15.3 (Podmínka prvního řádu). *Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je diferencovatelná. Funkce f je konvexní (na celém $\mathbb{R}^n$)*[3]*, právě když*

$$\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \implies f(\mathbf{y}) \geq f(\mathbf{x}) + f'(\mathbf{x})\,(\mathbf{y} - \mathbf{x}).$$

Věta 15.4 (Podmínka druhého řádu). *Nechť funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je dvakrát diferencovatelná. Funkce f je konvexní (na celém $\mathbb{R}^n$), právě když v každém bodě $\mathbf{x} \in \mathbb{R}^n$ je Hessova matice $f''(\mathbf{x})$ positivně semidefinitní.*

Nezáporné lineární kombinace konvexních funkcí jsou konvexní.
Složení konvexních funkcí gf je konvexní, pokud je funkce g nezáporná.
Maximum konvexních funkcí je také konvexní funkce.

Konvexní mnohostěny (polyedry)

Rádi kombinujeme vektory:
Vážený součet $\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k$ vektorů $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$ se nazývá jejich

lineární kombinace,	jestliže	$\alpha_1, \dots, \alpha_k \in \mathbb{R}$.		
afinní kombinace,	jestliže	$\alpha_1, \dots, \alpha_k \in \mathbb{R}$,	$\alpha_1 + \cdots + \alpha_k = 1$.	
nezáporná kombinace,	jestliže	$\alpha_1, \dots, \alpha_k \in \mathbb{R}$,		$\alpha_1, \dots, \alpha_k \geq 0$.
konvexní kombinace,	jestliže	$\alpha_1, \dots, \alpha_k \in \mathbb{R}$,	$\alpha_1 + \cdots + \alpha_k = 1$,	$\alpha_1, \dots, \alpha_k \geq 0$.

Množina, která je uzavřená vůči

lineárním kombinacím,	se nazývá	**lineární podprostor**.
afinním kombinacím,	se nazývá	**afinní podprostor**.
nezáporným kombinacím,	se nazývá	**konvexní kužel**.
konvexním kombinacím,	se nazývá	**konvexní množina**.

Uzavřený poloprostor:

$\{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} \geq b \,\}.$

Konvexní mnohostěn:

Konvexní mnohostěn (krátce jen **mnohostěn**, angl. *polyhedron*) je průnik konečně mnoha uzavřených poloprostorů. Je to tedy množina

$$X = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^T\mathbf{x} \geq b_i \ \ \forall i = 1, \dots, m \,\} = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b} \,\}, \qquad (12.5)$$

Dimenze mnohostěnu je dimenze jeho afinního obalu.

Extremální body:

Bod $\mathbf{x} \in X$ se nazývá **extremální bod** konvexní množiny X, jestliže neexistují dva různé body z X takové, že $\mathbf{x}$ je střed úsečky spojující tyto dva body, tj. jestliže platí implikace

$$\mathbf{x}_1, \mathbf{x}_2 \in X, \ \ \mathbf{x} = \tfrac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2) \implies \mathbf{x}_1 = \mathbf{x}_2. \qquad (12.6)$$

Počet extremálních bodů mnohostěnu může být **exponenciální** počtu nerovnic.

Pro neprázdný konvexní mnohostěn je ekvivalentní:
Mnohostěn má aspoň jeden extremální bod.
Mnohostěn neobsahuje přímku.

Věta 12.5. *Mějme konvexní mnohostěn, který neobsahuje přímku. Jestliže lineární funkce má na tomto mnohostěnu minimum, pak tato funkce nabývá na mnohostěnu minima aspoň v jednom z jeho extremálních bodů.*

Duální úloha

Konstrukce:

min	$\sum_{j\in J} c_j x_j$	max	$\sum_{i\in I} b_i y_i$	
za podm.	$\sum_{j\in J} a_{ij} x_j = b_i$	za podm.	$y_i \in \mathbb{R}$,	$i \in I_0$
	$\sum_{j\in J} a_{ij} x_j \geq b_i$		$y_i \geq 0$,	$i \in I_+$
	$\sum_{j\in J} a_{ij} x_j \leq b_i$		$y_i \leq 0$,	$i \in I_-$
	$x_j \in \mathbb{R}$		$\sum_{i\in I} a_{ij} y_i = c_j$,	$j \in J_0$
	$x_j \geq 0$		$\sum_{i\in I} a_{ij} y_i \leq c_j$,	$j \in J_+$
	$x_j \leq 0$		$\sum_{i\in I} a_{ij} y_i \geq c_j$,	$j \in J_-$

Schéma funguje na obě strany.

Občas se hodí matice:

min	$\mathbf{c}^T\mathbf{x}$	max	$\mathbf{b}^T\mathbf{y}$
za podm.	$\mathbf{A}\mathbf{x} \geq \mathbf{b}$	za podm.	$\mathbf{y} \geq \mathbf{0}$
	$\mathbf{x} \geq \mathbf{0}$		$\mathbf{A}^T\mathbf{y} \leq \mathbf{c}$

Věta 14.1 (o slabé dualitě). *Nechť $\mathbf{x}$ je přípustné primární řešení a $\mathbf{y}$ přípustné duální řešení. Pak $\mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\mathbf{y}$.*

Věta 14.4 (o silné dualitě). *Primární úloha má optimální řešení, právě když má duální úloha optimální řešení. Má-li primární úloha optimální řešení $\mathbf{x}$ a duální úloha optimální řešení $\mathbf{y}$, pak $\mathbf{c}^T\mathbf{x} = \mathbf{b}^T\mathbf{y}$.*

Věta 14.3 (o komplementaritě). *Nechť $\mathbf{x}$ je přípustné primární řešení a $\mathbf{y}$ přípustné duální řešení. Pak $\mathbf{c}^T\mathbf{x} = \mathbf{b}^T\mathbf{y}$ právě tehdy, když zároveň platí tyto dvě podmínky:*

$$\sum_{j\in J} a_{ij}x_j = b_i \quad \text{nebo} \quad y_i = 0 \qquad \forall i \in I, \qquad (14.4a)$$

$$x_j = 0 \quad \text{nebo} \quad \sum_{i\in I} a_{ij}y_j = c_j \qquad \forall j \in J. \qquad (14.4b)$$

Podmínky (14.4) budeme nazývat *podmínky komplementarity.* Říkají, že na každém řádku ve dvojici duálních úloh je vždy alespoň jedno omezení aktivní, buď primární nebo duální (přičemž omezení typu rovnosti považujeme vždy za aktivní).

Věta 14.5. *Z devíti možností pro dvojici duálních úloh se realizují tyto:*

primární/duální	*má optimum*	*neomezená*	*nepřípustná*
má optimum	*ano*	*ne*	*ne*
neomezená	*ne*	*ne*	*ano*
nepřípustná	*ne*	*ano*	*ano*