

# Naučno izračunavanje — Belekške

Andrija Urošević

## Rešavanje problema matematičkim metodama

- **Modelovanje**
  - Relevantne veličine, njihovo kvantitativno izražavanje i odnos između njih: *matematički model*  $M$ .
  - Pitanje na koje želimo dobiti odgovor: *matematički problem*  $P$ .
- **Rešavanje**
  - Primena metode koja može rešiti problem  $P$ .
  - Dobijamo rešenje  $S$ .
- **Interpretacija**
  - Rešenje  $S$  u modelu  $M$ , interpretirano u terminima polaznog problema.

## Modelovanje problema

- Poteškoće pri modelovanju
  - Potrebno je procizno uočiti relevantne vrednosti i odnose između njih, treba opisati odgovarajućim formalnim matematičkim jezikom.
  - Kako modelujemo problem direktno utiče na to koji metod rešavanja možemo da primenimo.
- Matematički model
  - Apstrakcija polaznog problema, kako se fokusiramo samo na relevantna svojstva problema.
  - Skup promenljivih predstavlja relevantne vrednosti.
  - Skup formula predstavlja relevantne odnose između tih vrednosti.
- Formulacija
  - Matematička teorija koja ima pogodna svojstva (diferencijabilnost, konveksnost, ...) se preporučuje pri formulisanju modela. Razlog tome je šira primena metoda za rešavanje tog problema.
- Pojednostavljenje modela
  - Model uprostimo sve dok je greška rešenja prihvatljiva.
  - Neke tehnike:
    - \* Zamena beskonačnih procesa konačnim
    - \* Zamena opštih matrica specifičnim matricama: blok dijagonalne, dijagonalne, trougaone, ...

- \* Zamena proizvoljnih funkcija jednostavnijim funkcijama: polinimima, konveksnim funkcijama. . . .
- \* Zamena nelinearnih problema linearnim problemima.
- \* Zamena diferencijalnih jednačina algebarskim jednačinam.
- \* Zamena beskonačno dimenzionih prostora konačno dimenzionim prostorima.
- Da bi tehnike pojednostavljenja bile relevantne potrebno je da:
  - \* alternativni problem možemo lakše rešiti, a čije rešenje nije drastično drugačije od polaznog;
  - \* transformacija tekućeg problema u lakši problem dozvoljava izračunavanje rešenja tekućeg problema pomoću rešenja lakšeg problema.
- Upozorenja:
  - Model ne oslikava precizno stvarnost
  - Model može biti dobar u nekim aspektima, a loš u drugim.
  - Podešavanje podataka dovodi do prilagođavanja modelu, u praksi ne daje dobre rezultate.
  - Ne treba se držati modela koji ne rade.

## Rešavanje problema

- Obično sam metod rešavanja dolazi na osnovu dobro izabranog modela problema.
- U nekim slučajevima sam model nema metodu koja može da se primeni.

## Interpretacija rešenja

- Kada dobijemo rešenje modela, primenjujemo inverzne transformacije pojednostavljivanja nad tim rešenjem.
- Transformisano rešenje razmatramo u terminima veza stvarnih fenomena i promenljivih u modelu.
  - Treba voditi računa o jedinicama.

## Aproksimacija i greške u izračunavanju

- Greške pre samog naučnog izračunavanja:
  - **Modelovanje:** Apstrakcija i pojednostavljenje dovode do greške
  - **Empirijska merenja:** Uključuju dozu neprekidnosti zbog nesavršenosti mernih instrumenata
  - **Prethodna izračunavanja:** Ulazni podaci mogu biti rezultat nekog prethodnog izračunavanja, pa se greška tako akumulira.
- Prethodni problemi nisu otkljivi, sledeća dva jesu:
  - **Diskretizacija i odsecanje:** Povećanjem granularnosti smanjujemo grešku. Beskonačne procese koje zamenjujemo konačnim možemo kontrolisati njihov broj koraka.
  - **Zaokruživanje:** Broj decimala koje se koriste za zapis realnih brojeva.

- Dve grupe grešaka: (1) *Greške podataka*; (2) *Greške izračunavanja*.
- **Procena greške.** Za pravu i približnu vrednost  $x$  i  $x'$  definišemo greške:
  - *Apsolutna greška*:  $E(x, x') = |x - x'|$ .
  - *Relativna greška*:  $R(x, x') = \frac{|x - x'|}{|x|}$

## Stabilnost, uslovljenost i regularizacija

- Algoritam je *nestabilan* ukoliko se njegova greška akumulira tokom njegovog izvršavanja, u suprotnom algoritam je *stabilan*.
- *Poništavanje* je slučaj kada je relativna greška mala usled oduzimanja realnih vrednosti koje nose grešku.
- Problem je *loše uslovljen* ako za malo različite podatke na ulozu daje drastično različita rešenja.
- Neka su  $\alpha$  ulazi podaci, i  $x(\alpha)$  rešenja problema  $P$ . Tada *uslovljenost* problem  $P$  definišemo kao

$$Cond(P) = \frac{R(x(\alpha), x(\alpha'))}{R(\alpha, \alpha')} = \frac{|x(\alpha) - x(\alpha')|/|x(\alpha)|}{|\alpha - \alpha'|/|\alpha|}.$$

- Uslovljenost funkcije  $f$ :

$$Cond(f) = \frac{|f(x) - f(x + \Delta x)|/|f(x)|}{|\Delta x|/|x|} \approx |xf'(x)/f(x)|$$

- Uslovljenost matrice  $A$ :

$$Cond(A) = |A^{-1}||A|$$

- Uslovljenost sistema  $Ax = b$ :

$$\begin{aligned} Cond(P) &= \frac{|A^{-1}b - A^{-1}(b + \Delta b)|/|A^{-1}b|}{|\Delta b|/|b|} \\ &= \frac{|A^{-1}\Delta b|/|A^{-1}b|}{|\Delta b|/|b|} \\ &= \frac{|A^{-1}\Delta b|}{|\Delta b|} \frac{|Ax|}{|x|} \end{aligned}$$

- Lošu uslovljenost rešavamo *regularizacijom*.
  - Zamenjujemo problem koji je loše uslovljen bliskim problemom koji je dobro uslovljen.
  - Razlika između ta dva problema treba da bude podesiva nekim parametrom, tj. kada parametar teži nuli problemi su jednaki.

## Aproksimacija funkcija

- Aproksimacija funkcije  $f$  je funkcija  $g$  koja je funkciji  $f$  bliska u nekom unapred definisanom smislu.

- Aproksimacija funkcija se vrši iz različitih razloga:
  - pojednostavljanje evaluacije funkcije;
  - zamenom funkcije nekom funkcijom sa boljim matematičkim osobinama;
  - ne znamo simboličku reprezentaciju funkcije već samo njene vrednosti u nekim tačkama.
- Postoje razni kriterijumi za aproksimaciju:
  - $\|f - g\|_2^2 = \int_a^b (f(x) - g(x))^2 dx$ ; (kriterijum je površina između dve funkcije)
  - $\|f - g\|_2^2 = \sum_{i=1}^n (f(x_i) - g(x_i))^2$ ; (ukupno odstupanje u svim tačkama u kojima je vrednost funkcije poznata)
  - $\|f - g\|_\infty = \sup_{x \in [a,b]} |f(x) - g(x)|$ . (samo najveće odstupanje je bitno)

## Primeri problema aproksimacije funkcija

- Problem linearne aproksimacije:
  - Aproksimacija:  $g(x, \alpha) = \alpha_0 + \sum_{i=1}^n \alpha_i x_i$ .
  - Kriterijum:  $\min_{\alpha} \sum_{i=1}^N (g(x_i, \alpha) - f(x_i))^2$ .
- Problem rekonstrukcije zamućene slike operatorom  $A$ :
  - $x = A^{-1}y$  ne daje dobro rešenje, kako je  $A$  loše uslovljena matrica.
  - Regularizacija obezbeđuje da se susedni pikseli ne razlikuju mnogo:

$$\min_x \|Ax - y\|^2 + \lambda \left( \sum_{i=1}^M \sum_{j=1}^{N-1} (x_{i,j} - x_{i,j+1})^2 + \sum_{i=1}^{M-1} \sum_{j=1}^N (x_{i,j} - x_{i+1,j})^2 \right)$$

- Problem konstrukcije slike od  $N$  slika različitih delova iste scene.
  - Moramo uračunati razlike među delovima slika: To su rotacija kamere za ugao  $\theta$ , translacija kamere za vektor  $(u, v)$  i skaliranje za vrednost  $s$ . Jedna takva veza može biti data matricom transformacije ( $a = s \cos \theta$ ,  $b = s \sin \theta$  i  $s = \sqrt{a^2 + b^2}$ ):

$$G = \begin{pmatrix} a & -b & u \\ b & a & v \\ 0 & 0 & 1 \end{pmatrix},$$

- Potrebno je još odrediti i upariti detalje na slikama (postoji algoritam). Neka je skup lokacija detalja  $\{x_{ij} | j = 1, \dots, M\}$ , i za svake dve slike  $i$  i  $j$  dat  $F(i, j)$  skup indeksa detalja koji su uspešno upareni.
- Konačan optimizacioni problem postaje ( $G_i$  matrica transformacije sa parametrima  $(a_i, b_i, u_i, v_i)$ ):

$$\min_{\mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{v}} \sum_{i=1}^N \sum_{j=i+1}^N \sum_{k \in F(i,j)} \|G_i x_{ik} - G_j x_{jk}\|^2$$

- Određivanje koordinata GPS uređaja:
  - $(u, v, w)$  koordinate GPS uređaja koje treba izračunati;
  - $(p_i, q_i, r_i)$  koordinate  $i$ -tog satelita;
  - $\rho_i$  udaljenost  $i$ -tog satelita od GPS uređaja.
  - Za svaki satelit treba da važi:  $\sqrt{(u - p_i)^2 + (v - q_i)^2 + (w - r_i)^2} = \rho_i$ .
  - Problem se svodi na:

$$\min_{u,v,w} \sum_{i=1}^n (\sqrt{(u - p_i)^2 + (v - q_i)^2 + (w - r_i)^2} - \rho_i)^2.$$

## Aproksimacija u Hilbertovim prostorima

- Vektorski prostor koji je kompletan u odnosu na metriku indukovanu skalarnim proizvodom  $d(x, y) = \|x - y\| = \sqrt{(x - y) \cdot (x - y)}$  se naziva *Hilbertovim prostorom*.
  - $\mathbb{R}^n$  je Hilbertov prostor
  - $\mathcal{L}_2[a, b]$  prostor funkcija koje su integrabile sa kvadratom na intervalu  $[a, b]$  je Hilbertov prostor.
- Sistem vektor  $\{e_i | i \in \mathbb{N}\}$  je *ortonormiran* ako

$$e_i \cdot e_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad i, j \in \mathbb{N}.$$

- Neka je  $\{e_i | i \in \mathbb{N}\}$  ortonormiran sistem vektora Hilbertovog prostora  $\mathcal{H}$ . Koeficijenti  $x \cdot e_i$  nazivaju se *Furijeovi koeficijenti* vektora  $x \in \mathcal{H}$ , a red  $\sum_{i=1}^{\infty} (x \cdot e_i) e_i$  se naziva *Furijevo red* vektora  $x \in \mathcal{H}$ .

**Teorema 1** Za ortonormirani sistem  $\{e_i | i \in \mathbb{N}\}$  u Hilbertovom prostoru  $\mathcal{H}$ , sledeća tvrđenja su ekvivalentna:

- Za svako  $x \in \mathcal{H}$  i svako  $\varepsilon > 0$ , postoje skalari  $\lambda_1, \lambda_2, \dots, \lambda_n$ , takvi da važi  $\|x - \sum_{i=1}^n \lambda_i e_i\| < \varepsilon$ .
- Za svako  $x \in \mathcal{H}$  važi  $\sum_{i=1}^{\infty} (x \cdot e_i) e_i = x$  (pri čemu se podrazumeva konvergencija u smislu metrike prostora  $\mathcal{H}$ )
- Za svako  $x \in \mathcal{H}$  važi  $\sum_{i=1}^{\infty} (x \cdot e_i)^2 = \|x\|^2$  (Parselova jednakost)
- Ako je vektor  $x \in \mathcal{H}$  takav da je  $x \cdot e_i = 0$  za svako  $i \in \mathbb{N}$ , onda važi  $x = 0$ .

**Teorema 2** Neka je  $f$  element Hilbertovog prostora  $\mathcal{H}$  i neka je  $\mathcal{H}'$  njegov potprostor čiju bazu čine elementi  $\{g_1, g_2, \dots, g_n\}$ . Postoji element najbolje aproksimacije  $g^* = \sum_{i=1}^n c_i^* g_i \in \mathcal{H}'$ , takav da važi

$$\left\| f - \sum_{i=1}^n c_i^* g_i \right\| = \inf_{c_1, \dots, c_n} \left\| f - \sum_{i=1}^n c_i g_i \right\|.$$

Dodatno, važi da je  $(f - g^*) \cdot x = 0$  za sve  $x \in \mathcal{H}'$  akko je  $g^*$  element najbolje aproksimacije za  $f$  iz  $\mathcal{H}'$ .

- Element najbolje aproksimacije za  $f$  je njegova ortogonalna projekcija na prostor  $\mathcal{H}'!!!$
- Koeficijenti najbolje aproksimacije se mogu odrediti iz sistema:

$$\sum_{i=1}^n c_i (g_i \cdot g_j) = f \cdot g_j, \quad j = 1, \dots, n$$

- Ako je baza  $\{g_1, \dots, g_n\}$  ortogonalna, svi skalarni proizvodi  $g_i \cdot g_j$  su jednaki nuli ako  $i \neq j$ , tako da u tom slučaju nije potrebno rešavati sistem jednačina već je dovoljno izračunati skalarne proizvode i izraziti koeficijente  $c_i$  iz dobijenih jednakosti u kojima učestvoje po jedan koeficijent  $c_i$ .

### Srednjevadratna aproksimacija

- Neka je  $\mathcal{L}_2[a, b]$  Hilbertov prostor funkcija integrabilnih sa kvadratom na intervalu  $[a, b]$ , u kome je norma definisana integralom  $\|f\|^2 = \int_a^b f^2(x) dx$  onda se element najbolje aproksimacije naziva elementom *najbolje srednjevadratne aproksimacije*.
- Ako je funkcija  $f$  definisana na konačnom skupu tačaka  $\{x_0, \dots, x_m\}$  integral zamenjujemo sumom, tj.  $\|f\|^2 = \sum_{i=1}^m f^2(x_i)$ .
- Metoda koja rešava srednjevadratnu aproksimaciju na konačnom skupu tačaka naziva se *metoda najmanjih kvadrata* (engl. *least squares method*).
- Sistem koji se rešava uzima sledeći oblik:

$$\sum_{i=1}^n c_i \sum_{k=1}^m g_i(x_k) g_j(x_k) = \sum_{k=1}^m f(x_k) g_j(x_k) \quad j = 1, \dots, n.$$

$$\sum_{k=1}^m g_j(x_k) \left( \sum_{i=1}^n c_i g_i(x_k) \right) = \sum_{k=1}^m f(x_k) g_j(x_k) \quad j = 1, \dots, n.$$

$$A^T A x = A^T b$$

$$x = (A^T A)^{-1} A^T b$$

- Prethodna jednačina predstavlja rešenje problema

$$\min_x \|Ax - b\|^2.$$

- Drugi način izvođenja rešenja:

$$\begin{aligned} \|Ax - b\|^2 &= (Ax - b)^T (Ax - b) \\ &= ((Ax)^T - b^T)(Ax - b) \\ &= (x^T A^T - b^T)(Ax - b) \\ &= b^T b - x^T A^T b - (b^T Ax)^{T^T} + x^T A^T Ax \\ &= b^T b - x^T A^T b - (x^T A^T b)^T + x^T A^T Ax \\ &= b^T b - 2x^T A^T b + x^T A^T Ax \end{aligned}$$

- Izjednačavanjem gradijenta po  $x$  sa nulom dobijamo:

$$2A^T Ax - 2A^T b = 0.$$

- Matrica  $(A^T A)^{-1} A^T$  je *Mur-Penrouzov pseudo inverz* matrice  $A$ .
- Metod srednjekvadratne aproksimacije se često koristi za rešavanje problema linearne regresije.

**Teorema 3 (Gaus-Markov)** Ukoliko važi  $E(\varepsilon) = 0$  i  $cov(\varepsilon) = \sigma^2$ , za konstantno  $\sigma^2 > 0$ , onda za ocenu  $\hat{w} = (X^T X)^{-1} X^T y$  važi

$$E(\hat{w}) = w, cov(\hat{w}) = \sigma^2 (X^T X)^{-1}.$$

Takođe, za svaku nepristrasnu linearnu ocenu  $\tilde{w}$  parametra  $w$  važi

$$\sum_{i=1}^n (w_i - \hat{w}_i)^2 \leq \sum_{i=1}^n (w_i - \tilde{w}_i)^2$$

- Ukoliko je matrica  $A^T A$  loše uslovljena (kolone ili vrste matrice  $A$  su visoko korelisane), tada se koristi regularizacija i rešava se problem (*Tihonovljeva regularizacija* ili *grebena regularizacija*):

$$\min_x \|Ax - b\|^2 + \lambda \|x\|^2$$

- Slično kao u prethodnom slučaju:

$$\|Ax - b\|^2 + \lambda \|x\|^2 = (Ax - b)^T (Ax - b) + \lambda x^T x = b^T b - 2x^T A^T b + x^T A^T A x + \lambda x^T x.$$

- Računanjem gradijenta po  $x$  i izjednačavanjem sa nulom dobijamo:

$$A^T Ax - A^T b + \lambda x = 0$$

$$x = (A^T A + \lambda I)^{-1} A^T b$$

- Uklanjanje šuma iz signala:

$$\min_x \|x - y\|^2 + \lambda \sum_{i=1}^{n-1} (x_i - x_{i+1})^2$$

– Uvodimo matricu  $D$ :

$$\begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & -1 \end{pmatrix}$$

– Dati problem postaje:

$$\min_x \|Ix - y\|^2 + \|\sqrt{\lambda} Dx - 0\|^2$$

$$\min_x \left\| \begin{pmatrix} I \\ \sqrt{\lambda} D \end{pmatrix} x - \begin{pmatrix} y \\ 0 \end{pmatrix} \right\|^2$$

– Odgovarajuće rešenje:

$$x = (I + \lambda D^T D)^{-1} y$$

- Rekonstrukcija zamućene slike:

$$\min_x \|Ax - y\|^2 + \lambda \left( \sum_{i=1}^M \sum_{j=1}^{N-1} (x_{i,j} - x_{i,j+1})^2 + \sum_{i=1}^{M-1} \sum_{j=1}^N (x_{i,j} - x_{i+1,j})^2 \right)$$

– Uvodimo matricu  $D_h$  i  $D_v$ :

$$\begin{pmatrix} I & -I & 0 & \dots & 0 & 0 & 0 \\ 0 & I & -I & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & I & -I & 0 \\ 0 & 0 & 0 & \dots & 0 & I & -I \end{pmatrix} \begin{pmatrix} D & 0 & \dots & 0 \\ 0 & D & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & D \end{pmatrix}$$

– Problem možemo zapisati kao:

$$\min_x \|Ax - y\|^2 + \|\sqrt{\lambda} D_v x - 0\|^2 + \|\sqrt{\lambda} D_h x - 0\|^2$$

$$\min_x \left\| \begin{pmatrix} A \\ \sqrt{\lambda} D_v \\ \sqrt{\lambda} D_h \end{pmatrix} x - \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix} \right\|^2$$

– Odgovarajuće rešenje:

$$x = (A^T A + \lambda D_v^T D_v + \lambda D_h^T D_h)^{-1} A^T y$$

## Furijeova transformacija

- Trigonometrijski Furijeov red se zasniva na sistemu različitih frekvencija  $\cos(kx)$  i  $\sin(kx)$  za  $k = 0, 1, \dots$
- Furijeovi koeficijenti omogućavaju analizu signala u odnosu na frekvencije koje su u njemu zastupljene, odnosno *spektar signala*. ???
- *Furijeova transformacija* prevodi reprezentaciju funkcije iz vremenskog domena u frekvencijski domen.
  - *Inverzna Furijeova transformacija* radi obrnuto.
  - Neke vrste Furijeovih transformacija: *razvoj u Furijeov red*, *neprekidna Furijeova transformacija* i *diskretna Furijeova transformacija*.
- Neka je funkcija  $f$  periodična i integrabilna na intervalu  $[a, b]$ . Tada se može razviti u Furijeov red:

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left( a_k \cos\left(\frac{2\pi kt}{b-a}\right) + b_k \sin\left(\frac{2\pi kt}{b-a}\right) \right),$$



$$a_k = \frac{2}{b-a} \int_a^b f(t) \cos\left(\frac{2\pi kt}{b-a}\right) dt, \quad k = 0, 1, 2, \dots$$

$$b_k = \frac{2}{b-a} \int_a^b f(t) \sin\left(\frac{2\pi kt}{b-a}\right) dt, \quad k = 1, 2, 3, \dots$$

- *Primer:*  $f(t) = 5 \cos(2t) + 3 \sin(8t)$  je periodična na intervalu  $[0, \pi]$ . Odatle svi Furijeovim koeficijenti su 0 sem  $a_1 = 5$  i  $b_4 = 3$ .
- Komleksna reprezentacije Furijeovog reda:

$$f(t) = \sum_{k=-\infty}^{\infty} \hat{f}_k e^{\frac{-2\pi i k t}{b-a}}$$

$$\hat{f}_k = \frac{1}{b-a} \int_a^b f(t) e^{\frac{2\pi i k t}{b-a}} dt$$

- Odnos između realne i kompleksne reprezentacije su u tesnoj vezi:

$$a_0 = 2\hat{f}_0$$

$$a_k = \hat{f}_k + \hat{f}_{-k}$$

$$b_k = i\hat{f}_{-k} - \hat{f}_k$$

- Za koeficijente važi  $\overline{\hat{f}_k} = \hat{f}_{-k}$ :

$$\begin{aligned} \overline{\hat{f}_k} &= \overline{\frac{1}{b-a} \int_a^b f(t) e^{\frac{2\pi i k t}{b-a}} dt} \\ &= \frac{1}{b-a} \int_a^b \overline{f(t) e^{\frac{2\pi i k t}{b-a}}} dt \\ &= \frac{1}{b-a} \int_a^b f(t) e^{\frac{-2\pi i k t}{b-a}} dt \\ &= \hat{f}_k \end{aligned}$$

- Promenljiva  $t$  predstavlja vreme, dok Furijeovi koeficijenti  $\hat{f}_k$  predstavljaju intenzitet odgovarajućih frekvencija u signalu.
  - U razvoju u Furijeov red vremenski domen je neprekidno, ali je frekvencijski domen diskretan, tj. periodična funkcija se može predstaviti preko beskonačno mnogo broja sinusa i kosinusa, ali sa diskretnim frekvencijama.
  - Ovaj problem se prevazilazi prelaskom sa reda na integral (Furijeova transformacija i inverzna Furijeova transformacija):

$$\hat{f}(u) = \int_{-\infty}^{+\infty} f(t) e^{2\pi i u t} dt$$

$$f(t) = \int_{-\infty}^{+\infty} \hat{f}(u) e^{-2\pi i u t} du$$

- Mana ovih metoda je što je funkcija  $f$  obično poznata samo na konačnom skupu tačaka.
- Neka su vrednosti funkcije  $f_j = f(t_j)$ , gde je  $t_j = t_0 + jh$ , za  $j = 0, 1, \dots, n-1$  i  $h > 0$ . Tada:

$$\hat{f}_k = \frac{1}{n} \sum_{j=0}^{n-1} f_j e^{\frac{k 2\pi i j}{n}} \quad k = 0, 1, \dots, n-1$$

$$f_k = \frac{1}{n} \sum_{j=0}^{n-1} \hat{f}_j e^{-\frac{2\pi i k j}{n}} \quad j = 0, 1, \dots, n-1$$

- *Primer* uklanjanje šuma: Signal  $\rightarrow$  (FFT)  $\rightarrow$  Frekvencije  $\rightarrow$  (clamp)  $\rightarrow$  Frekvencije (bez visokih)  $\rightarrow$  (IFFT)  $\rightarrow$  Signal (bez šuma).
- Furijeova transformacija u dve dimenzije:
  - Nekretna:

$$\hat{f}(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{2\pi i (xu + yv)} dx dy$$

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \hat{f}(u, v) e^{-2\pi i (xu + yv)} du dv$$

- Diskretna:

$$\hat{f}_{lm} = \frac{1}{PQ} \sum_{j=0}^{P-1} \sum_{k=0}^{Q-1} f_{jk} e^{2\pi i \left( \frac{j l}{P} + \frac{k m}{Q} \right)}$$

$$f_{jk} = \sum_{l=0}^{P-1} \sum_{m=0}^{Q-1} \hat{f}_{lm} e^{-2\pi i \left( \frac{j l}{P} + \frac{k m}{Q} \right)}$$

- Koeficijenti Furijeove transformacije su, kao kompleksni brojevi, određeni modulom ili *amplitudom* i argumentom ili *fazom*.
  - Amplituda predstavlja jačinu nekog signala.
  - Faza predstavlja pomeraj frekvencije duž vremenske ose.
- *Dirakova delta funkcija* i  $f(x, y) = \delta(x, y) = \delta(x)\delta(y)$
- *Odsecanje dela spektra*:
  - Odsecanje viših frekvencija omogućava grub prikaz slike (uklanja ivice)
  - Odsecanje nižih frekvencija omogućava prepoznavanje ivica (istače ivice)
  - Uklanjanjem prepoznatljivih maksimuma uklanjaju se poreiodične strukture na slici.

### Brza Furijeova transformacija

- DFT (Diskretna Furijeova transformacija) ima složenost  $\Theta(n^2)$ .
- FFT (Brza Furijeova transformacija) ima složenost  $\Theta(n \log n)$ .
- Uvodimo  $n$ -ti koren jedinice  $w = e^{\frac{2\pi i}{n}}$ ,
  - Važi  $w^n = 1$ :

$$w^n = e^{\frac{2\pi i}{n}n} = e^{2\pi i} = 1$$

- Važi  $w^{k+\frac{n}{2}} = -w^k$ :

$$w^{k+\frac{n}{2}} = e^{\frac{2\pi i}{n}(k+\frac{n}{2})} = e^{\frac{2\pi i k}{n} + \pi i} = e^{\pi i} e^{\frac{2\pi i k}{n}} = -w^k$$

- Diskretna Furijeovra transformacija postaje:

$$\hat{f}_k = \frac{1}{n} \sum_{j=0}^{n-1} f_j w^{kj} \quad k = 0, 1, 2, \dots, n-1$$

- Važi  $\hat{f}_{k+n} = \hat{f}_k$ :

$$\hat{f}_{k+n} = \frac{1}{n} \sum_{j=0}^{n-1} f_j w^{(k+n)j} = \frac{1}{n} \sum_{j=0}^{n-1} f_j w^{kj} = \hat{f}_k$$

- Koeficijenti se mogu izračunati preko parnih i neparnih elementa:

$$\begin{aligned} \hat{f}_k &= \frac{1}{n} \sum_{j=0}^{n-1} f_j w^{kj} = \frac{1}{n} \sum_{j=0}^{n/2-1} f_{2j} w^{2jk} + \frac{1}{n} \sum_{j=0}^{n/2-1} f_{2j+1} w^{(2j+1)k} = \\ &= \frac{1}{2} \frac{1}{n/2} \sum_{j=0}^{n/2-1} f_{2j} w^{2jk} + \frac{1}{2} w^k \frac{1}{n/2} \sum_{j=0}^{n/2-1} f_{2j+1} w^{(2j+1)k} = \frac{1}{2} (E_k + w^k O_k) \end{aligned}$$

- Takođe, važi:

$$E_{k+n/2} = E_k$$

$$O_{k+n/2} = O_k$$

- Dobijamo:

$$\hat{f}_k = \begin{cases} \frac{1}{2} (E_k + w^k O_k) & 0 \leq k < \frac{n}{2} \\ \frac{1}{2} (E_{k-\frac{n}{2}} + w^k O_{k-\frac{n}{2}}) & \frac{n}{2} \leq k < n \end{cases}$$

- Konačno:

$$\hat{f}_k = \frac{1}{2} (E_k + w^k O_k)$$

$$\hat{f}_{k+\frac{n}{2}} = \frac{1}{2} (E_k - w^k O_k)$$

- Algoritam FFT se može koristiti i kao algoritam za inverzni FFT, tako što se pre primene algoritma FFT, ulaz konjuguje, a nakon primene algoritma FFT, izlaz konjuguje.

## Konvolucija

- Da li postoji neka aritmetička veza između operacija nad signalima i nekih operacija nad njihovim Furijeovim transformacijama?

– Važi:

$$\begin{aligned}
 \widehat{f+g}(u) &= \int_{-\infty}^{+\infty} (f+g)(t)e^{2\pi iut} dt \\
 &= \int_{-\infty}^{+\infty} (f(t) + g(t))e^{2\pi iut} dt \\
 &= \int_{-\infty}^{+\infty} f(t)e^{2\pi iut} dt + \int_{-\infty}^{+\infty} g(t)e^{2\pi iut} dt \\
 &= \hat{f}(u) + \hat{g}(u)
 \end{aligned}$$

– Takođe,

$$\begin{aligned}
 \hat{f}(u)\hat{g}(u) &= \int_{-\infty}^{+\infty} f(x)e^{2\pi iux} dx \int_{-\infty}^{+\infty} g(y)e^{2\pi iuy} dy \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x)g(y)e^{2\pi i(x+y)u} dx dy \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x)g(v-x)e^{2\pi ivu} dx dv \\
 &= \int_{-\infty}^{+\infty} \left( \int_{-\infty}^{+\infty} f(x)g(v-x) dx \right) e^{2\pi ivu} dv \\
 &= \int_{-\infty}^{+\infty} (f * g)(v)e^{2\pi ivu} dv
 \end{aligned}$$

$$(f * g)(v) = \int_{-\infty}^{+\infty} f(x)g(v-x) dx$$

– Operacija  $*$  se naziva operacijom *konvolucije*.

**Teorema 5 (o konvoluciji):**

$$\widehat{f * g} = \hat{f}\hat{g}$$

$$\widehat{fg} = \hat{f} * \hat{g}$$

$$f * g = g * f$$

$$(f * g) * h = f * (g * h)$$

$$f * (g + h) = f * g + f * h$$

$$f * \delta = f$$

- Konvolucija u diskretnom smislu:

$$(f * g)_i = \sum_{j=0}^{n-1} f_j g_{i-j} \quad i = 0, 1, \dots, n-1$$

- Konvolucija u dve dimenzije:

$$(f * g)(u, v) = \int_{-\infty}^{+\infty} f(x, y) g(u - x, v - y) dx dy$$

$$(f * g)_{i,j} = \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} f_{k,l} g_{i-k,j-l}$$

- Po definiciji konvolucija dva signala ima vremensku složenost  $\Theta(n^2)$ . Ali primenom FFT algoritma, konvoluciju možemo izračunati u  $\Theta(n \log n)$  (zbog teoreme o konvoluciji  $\widehat{f * g} = \hat{f} \hat{g}$ ).
  - $\widehat{f * g} \text{ (FFT)} \rightarrow \hat{f} \hat{g} \text{ (množenje)} \rightarrow \hat{f} \hat{g} \text{ (teorema o konvoluciji)} \rightarrow \widehat{f * g} \text{ (IFFT)} \rightarrow f * g$
- *Primer:* Množenje polinoma je konvolucija

$$f(x) = \sum_{i=1}^m f_i x^i \quad g(x) = \sum_{i=1}^n g_i x^i$$

- Proizvod je veličine  $m + n + 1$ , te ulazne podatke proširujemo  $(f_1, f_2, \dots, f_m, 0, \dots, 0)$  i  $(g_1, g_2, \dots, g_n, 0, \dots, 0)$ .
- Množenje polinoma ima složenost  $\Theta((n + m) \log(n + m))$ .
- Konvolucija se koristi tako što je jedna funkcija signal, a druga funkcija predstavlja neku jednostavnu funkciju kojom transformišemo signal. Tu funkciju zovemo *filter*.
- Filter Gausovog zamučivanja.
  - Gausovo zvono:  $\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$
  - Uprošćeni filter zamučivanja:

$$\frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

- Filteri za otkrivanje ivica:
  - Sobel-Feildmonove vertikalne i horizontalne ivice:

$$G_x = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} * A \quad G_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} * A$$

- Aproksimacija intenziteta gradijenta je onda  $G = \sqrt{G_x^2 + G_y^2}$
- Brzo pronalaženje uzorka slike  $g$  u drugoj slici  $f$ .

## Osnovni koncept obrade signala

### Uzorkovanje

- Zbog prirode operisanja računara signal se opisuje diskretnim reprezentacijama.
  - Procedura uzorkovanja signala se radi tako što se odaberu vremenski trenuci u kojima će se meriti jačina zvuka, a kvantizaija odabir numeričke skale za predstavljanje izmerenih vrednosti.
  - Ako se na svakih  $T$  sekundi vrši uzorkovanje signala, govori se o uzorkovanju signala sa frekvencijama uzorkovanja  $f_s$ .
  - Veza između učestalosti uzorkovanja i frekvencije uzorkovanja:

$$f_s = \frac{1}{T};$$

$$w_s = \frac{2\pi}{T} = 2\pi f_s.$$

- Ukoliko postoji neka frekvencija  $f_b$  za koju važi  $f_b > f$  onda je frekvencija  $f$  ograničena frekvencijom  $f_b$  koju još i nazivamo granična frekvenija.
- *Najkvistov teorema*: Signal se može verodostojno reprodukovati samo ako je frekvencija uzorkovanja više od dva puta veća od granične frekvencije.
  - Kako ljudsko uvo čuje do oko  $22KHz$ , najčešće se vrši uzorkovanje od  $44.1KHz$ .
  - Kod video zapisa uzorkovanje se mora vršiti i više od dva puta učestalije, jer može pokazivati statična kretanja, pa čak i kretanja unazad.

### Curenje spektra

- Razvoj u Furijeov red i diskretna Furijeova transformacija pretpostavljaju periodičnost signala i diskretni frekvencijski domen, što u realnosti obično nije slučaj.
  - Čak iako je signal periodičan to nam ne garantuje da će njegovo uzorkovanje biti periodično.
    - \* To stvara skokove u vremenskom domenu i odgovarajuće visoke frekvencije u frekvencijskom domenu.
  - Ako kalibrišemo softver za analizu spektra tako da izražava frekvencije u celobrojnim Herima, onda:
    - \* Signal frekvencije  $3Hz$  prilikom Furijeove transformacije daje vrlo jasan pik na frekvencije  $3Hz$ ;
    - \* Signal frekvencije  $2.8Hz$  prilikom Furijeove transformacije će se predstaviti u celom frekvencijskom spektru. Ovaj fenomen se naziva *curenje spektra*.
- Problem se ublažuje pomoću *prozorskih funkcije*, tako što se signal u vremenskom domenu množi nekom prozorskom funkcijom.
  - Osobine prozorskih funkcija:
    - \* svuda izvan intervala su nula;
    - \* na krajevima intervala teže nuli;

- \* maksimum dostižu na sredini intervala.
- Neke prozorske funkcije:
  - Blekmenova
  - Hanova
  - Hamingova funkcija
  - Kasijerova funkcija

### Filtriranje signala

- Filtriranje signala obično se koristi kao prvi korak u selekciji informacija koje signal nosi.
  - Prvi način: FFT  $\rightarrow$  modifikacija signala  $\rightarrow$  IFFT
  - Drugi način: Konvolucija gde je jedan signal polazni signal, a drugi je filter.
  - Zašto je drugi pristup brži? Jer filteri obično vrlo malo ne nula vrednosti.
- *Linearni vremenski invarijantni sistemi* koji za linearne kombinacije ulaznih signala generišu linearne kombinacije izlaznih signala i čije ponašanje se ne menja u zavisnosti od vremena.
  - Linearno invarijentni sistem  $H$  koji preslikava signal  $x(t)$  u  $y(t)$  opisuje se kao:
    - \*  $H(ax(t)) = aH(x(t))$
    - \*  $H((x_1 + x_2)(t)) = H(x_1(t)) + H(x_2(t))$
  - $y(t) = tx(t)$  nije vremenski invarijantan
  - $y(t) = 2x(t)$  jeste vremenski invarijantan
- *Impulsni odgovor sistema* predstavlja kratkotrajni signal u vremenskom domenu, koji je najčešće diskretna Dirakova  $\delta$  funkcija.
  - Formalno, ako je ulaz sistema signal  $x[n] = \delta[n]$ , impulsni odgovor sistema je signal  $h[n] = H(\delta[n])$ .
- Diskretni signali zajedno sa odgovarajućim izlazima se mogu zapisati

$$x[n] = \sum_k x[k]\delta[n - k]$$

$$y[n] = H(x[n]) = \sum_k x[k]h[n - k]$$

- Podela filtera:
  - Po prirodi signala: analogni ili digitalni
  - Po dužini impulsnog odgovora: *filteri sa konačnim trajanjem impulsnog odgovora* (FIR filteri) i *filteri sa beskonačnom dužinom impulsnog odgovora* (IIR filteri)
- FIR filteri se mogu predstaviti kao konačne težinske sume prethodnih, trenutnih, ili budućih ulaza:

$$y[n] = \sum_{i=-M_1}^{M_2} b_i x[n - i]$$

- Primer:  $y[n] = \frac{1}{3}(x[n] + x[n-1] + x[n-2])$
- Frekvencijski odgovor sistema oslikava kako sistem reaguje na ulaze (harmonike)

$$x[n] = e^{i\omega n}$$

$$y[n] = \sum_{k=0}^M b_k x[n-k] = \sum_{k=0}^M b_k e^{i\omega(n-k)} = e^{i\omega n} \sum_{k=0}^M b_k e^{-i\omega k}.$$

- Frekvencijski odgovor filtera, podrazumeva:

$$H(\omega) = \sum_{k=0}^M b_k e^{-i\omega k}$$

- Kako se signali mogu razložiti na harmonike, dovoljno je poznavati frekvencijski odgovor filtera da bi filter bio definisan.
  - Ukoliko je amplituda frekvencijskog odgovora filtera za harmonik neke frekvencije jednaka nuli, to znači da taj filter eliminiše tu frekvenciju iz signala.
- Zamućenje prostim uprosečavanjem, Gausovo zamućenje i Sobel-Feldmanovi filteri predstavljaju FIR filtere.
- IIR filteri zavse od tekućih ulaza, prethodnih ulaza, i prethodnih izlaza:

$$y[n] = \sum_{l=1}^N a_l y[n-l] + \sum_{k=0}^M b_k x[n-k]$$

- Primer:  $y[n] = a_1 y[n-1] + b_0 x[n]$ 
  - \* Impulsni odgovor određujemo zamenom  $x[n] = \delta[n]$ .
  - \* Pretpostavljamo da važi  $x[0] = 0$  i  $y[0] = 0$ .
  - \* Rekurentno stižemo do rešenja:

$$y[n] = a_1^n b_0 x[0]$$

- Impulni odgovor jednog IIR filtera može biti:
  - Low-pass, High-pass, Band-pass, Band-stop

## Talasići

- Sistem trigonometrijskih funkcije nije uvek najbolji izbor.
  - Nije pogodan za funkcije koje nisu periodične.
  - Ne dozvoljava lokalizaciju: Ako je u nekom frekvencija prisutna u signalu, biće prisutna takom celog trajanja signala (u intervalima u kojima nije izražena biće poništena drugim frekvencijama).
  - Nije pogodan za funkcije koje nisu glatke.
- Definiše se funkcija koja ne mora biti glatka, pa čak ni neprekidna koju nazivamo *talasićem*. Ona generiše sistem talasića translacijama i skaliranjem.



- Primer osnovnog talasića je Harova funkcija:

$$\phi(x) = \begin{cases} 1 & x \in [0, \frac{1}{2}) \\ -1 & x \in [\frac{1}{2}, 1) \\ 0 & x \notin [0, 1) \end{cases}$$

- Ortonormirani sistem kojim se mogu proizvoljno dobo aproksimirati funkcije prostora  $L^2(\mathbb{R})$  su definisani kao:

$$\phi_{ij} = 2^{i/2} \phi(2^i x - j) \quad i, j \in \mathbb{Z}$$

- Mogu se koristiti i drugi osnovni talasići, ali je bitno od njih konstruisati ortonormirani sistem.

## Numerička linearna algebra

- Problemi rešavanja sistema jednačina, inverzija matrica, dokompozicija matrica, izračunavanje sopstevnih vektora i sopstvenih vrednosti matrica.

### Primeri problema numeričke linearne algebre

- Problem rangiranja stranica na internetu (*PageRank* algoritam).
- Internet se modeluje usmerenim grafom.
- Skup čvorova  $S$  je skup strana,  $|S| = N$ . Neka je  $K$  skup strana bez veza ka drugim stranama.
- Verovatnoća pristupa nekoj stranici  $s \in S$ :

$$P(s) = \sum_{r \in N(s)} \frac{P(r)}{n_r} + \sum_{r \in K} v_s P(r)$$

- $N(s)$  je skup strana koje pokazuju na stranu  $s$ .
- $n_r$  broj strana na koje se pokazuje sa strane  $r$ .
- $v = (v_1, \dots, v_s, \dots, v_N)$  vektor preference korisnika u posećivanju određene stranice.
- $p = (p_1, \dots, p_s, \dots, p_N)$  vektor verovatnoće posete svih stranica.
- $A$  matrica takva da  $A_{ij} = \frac{1}{n_j}$  ukoliko postoji veza sa strane  $j$  ka strani  $i$ , a 0 u suprotnom.
- $k = (k_1, \dots, k_s, \dots, k_N)$  vektor indikator da li čvor  $i$  nema naslednika.

$$p = (A + vk^T)p$$

- $p$  je sopstveni vektor matrice  $A + vk^T$ , pa se problem svodi na pronalaženje sopstvenog vektora matrice.
  - Matrica  $A + vk^T$  je stohastička, pa je njena najveća sopstvena vrednost jednaka 1, što znači da će sopstveni vektor imati odgovarajuću sopstvenu vrednost 1.

\* Ovo ne znači da postoji samo jedan sopstveni vektor koji ima sopstvenu vrednost 1.

- Korisnik ne bira uvek samo stranice dostupne sa tekuće strane već nekada bira narednu stranu nezavisno od tekuće:

$$p = \alpha(A + vk^T)p + (1 - \alpha)v = (\alpha(A + vk^T) + (1 - \alpha)ve^T)p$$

- $e$  vektor čiji su svi elementi jednaki 1, pri čemu poslednja jednakost važi zahvaljujući tome što važi  $e^T p = \sum_{i=1}^n p_i = 1$ .
- $p$  je sopstveni vrednost google matrice:

$$G = \alpha(A + vk^T) + (1 - \alpha)ve^T$$

- $p$  je sopstveni vektor koji odgovara najvećoj sopstvenoj vrednosti matrice  $G$ , ona koja je jednaka 1.
- *Norma* je funkcija  $\|\cdot\| : X \mapsto \mathbb{R}$  takva da za svako  $\alpha \in \mathbb{R}$  i  $x, y \in X$  važi:
  - $\|\alpha x\| = |\alpha| \|x\|$
  - $\|x + y\| \leq \|x\| + \|y\|$
  - Ako važi  $\|x\| = 0$ , onda važi  $x = 0$ .
- $p$ -norme nad vektorima iz  $\mathbb{R}^n$ :

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^n x_i^p}$$

- *Ekvivalentnost  $p$ -normi*: Za svake dve  $p$ -norme  $\|\cdot\|_a$  i  $\|\cdot\|_b$  postoje konstante  $0 < c_1 \leq c_2$ , takve da za svako  $x \in X$  važi:

$$c_1 \|x\|_b \leq \|x\|_a \leq c_2 \|x\|_b$$

– Konvergencija u jednoj  $p$ -normi povlači konvergenciju u drugoj  $p$ -normi.

- $p$ -norma nad matricama:

$$\|A\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

- Frobenijusova norma:

$$\|A\|_F^2 = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$$

## Dekompozicija matrica

- LU, Čoleski, QR i SVD dekompozicije

## LU dekompozicija

- LU dekompozicija podrazumeva predstavljanje matrice u vidu proizvoda  $L$  donjetrougaone ja jedinicama na dijagonali, i  $U$  gornjetrougaone.

$$A = LU$$

- Koristimo za rešavanje sistema:  $Ax = b$

$$Ax = (LU)x = L(Ux) = b$$

- Prvo rešimo:  $Ly = b$
- Drugo rešimo:  $Ux = y$
- LU dekompozicija zahteva  $\frac{1}{3}n^3$  množenja, dok Gaus-Žordanova metoda zahteva  $n^3$  množenja.
- Rešavanje sistema:  $Ax = b_1, \dots, Ax = b_n$  je značajno ubrzano.
- Inverzna matrica  $A^{-1}$ :
  - Prvo rešimo:  $Ly = I_i$
  - Drugo rešimo:  $Ux = y$
- Pod čvrstim uslovima moguće je dobiti LU dokompiziciju matrice
  - Ako to nije moguće uvek je moguće razmeniti njene redove i dobiti odgovarajuću dekompoziciju.
- Ako je poznata LU dekompozicija, lako je izračunati determinantu matrice (proizvod dijagonalnih elemenata matrice  $U$ ).
  - Treba voditi računa o broju permutacija, i tako odrediti znak determinante.

## Čoleski dekompozicija

- Matrica je pozitivno definitna ukoliko za sve vektore  $x$  odgovarajuće dimenzije, različite od nule, važi  $x^T Ax > 0$ .
- Čoleski dekompoziciju kvadratne matrice  $A$  je moguće izvršiti pod uslovima da je ona simetrična i pozitivno definitna ( $L$  je donje trougaona matrica sa strogo pozitivnim dijagonalnim elementima):

$$A = LL^T$$

- Matricu  $L$  računamo narednim formulama:

$$l_{ii} = \left( a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right) \quad i = 1, \dots, n$$

$$l_{ji} = \frac{1}{l_{ii}} \left( a_{ij} - \sum_{k=1}^{i-1} l_{ik} l_{jk} \right) \quad i = 1, \dots, n-1, j = i+1, \dots, n$$

- Izračunavanje Čoleski dekompozicije zahteva  $\frac{1}{6}n^3$  množenja, što je duplo bolje od  $LU$  dekompozicije.

- Pozitivna definitnost implicira invertibilnost, pa je stabilost izračunavanja veća.
- Ako je poznata Čoleski dekompozicija matrice  $A$  rešavanje sistema  $Ax = b$  se vrši tako što se redom reše sistemi:

$$Ly = b$$

$$L^T x = y$$

- Determinantu je moguće izračunati na sledeći način:

$$\det(A) = \det(L) \cdot \det(L^T) = (\det(L))^2$$

- $\det(L)$  je proizvod njenih dijagonalnih elemenata.
- Postojanje Čoleski dekompozicije matrice implicira da je ta matrica simetrična i pozitivno definitna.
  - Simetrija je očigledna
  - Pozitivno semidefinitna:  $x^T Ax = x^T LL^T x = (L^T x)^T (L^T x) = \|L^T x\|^2 \geq 0$
  - Pozitivno definitna:  $x^T Ax = 0$  akko  $x = 0$ .
- Sprovođenjem Čoleski dekompozicije možemo utvrditi da li je matrica pozitivno definitna.

## QR dekompozicija

- Za kvadratnu matricu  $Q$  se kaže da je ortogonalna, ukoliko za nju važi  $Q^T Q = Q Q^T = I$ 
  - Nije potrebno nalaziti inverz ovakvih matrica.
  - Čuvaju euklidsku normu vektora pri množenju:

$$\|Qx\|_2^2 = (Qx)^T Qx = x^T Q^T Qx = x^T x = \|x\|_2^2$$

- Uslovljenost matrice se definiše kao maksimum uslovljenosti sistema jedinačin  $Ax = b$ :

$$\max_{x, \Delta b} \frac{\|Q^{-1} \Delta b\|}{\|\Delta b\|} \cdot \frac{\|Qx\|}{\|x\|} = \max_{x, \Delta b} \frac{\|Q^T \Delta b\|}{\|\Delta b\|} \cdot \frac{\|Qx\|}{\|x\|} = 1$$

- Ortogonalne matrice imaju minimalnu moguću uslovljenost  $Cond(Q) = 1$
- Proizvoljna matrica  $A$ , dimenzije  $m \times n$  se može predstaviti pomoću ortogonalne matrice  $Q$  dimenzije  $m \times m$  i matrice  $R$  dimenzije  $m \times n$  u obliku

$$A = QR$$

- Matrice  $R$  je oblika  $R = \begin{pmatrix} R' \\ 0 \end{pmatrix}$ , gde je  $R'$  gornjetrogona matrica dimen-

zija  $n \times n$ . Matrica  $R$  ima formu

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

- Zbog nula u matrici  $R$  moguće je čuvati samo njenu redukovanu formu  $R'$  i odgovarajuću redukovanu formu  $Q'$ .
- Ako je matrica  $A$  kvadratna, sistem jednačina  $Ax = b$  se može rešiti kao ( $Q^{-1} = Q^T$ ):

$$Rx = Q^T b$$

- QR dekompozicija zahteva više operacija nego izračunavanje LU dekompozicije.
- Matrica  $R$  je gornjetrougaona, a matrica  $Q$  ortogonalna, pa QR dekompozicija može voditi boljoj numeričkoj stabilnosti
- Posebno je interesantno koristi QR dekompoziciji kada matrica  $A$  nije kvadratna, što prethodne dve dekompozicije zahtevaju da je matrica  $A$  kvadratna.
- Neka je  $A = QR$ , prostor kolona matrice  $A$  je prostor kolona matrice  $Q$ , tj.

$$\begin{aligned} u \in C(A) &\iff u = Ax \text{ za neko } x \\ &\iff u = QRx \\ &\iff u = Qy \text{ za } y = Rx \text{ (matrica } Q \text{ je ortogonalna pa je } A \text{ invertibilna)} \\ &\iff u \in C(Q) \end{aligned}$$

- Algoritam koji se koristi za računanje QR dekompozicije je Householderov algoritam, koji je numerički stabilan, zahvaljujući tome što počina na upotrebi ortogonalnih matrica, a zahteva oko  $2mn^2 - \frac{2}{3}n^3$  operacija.
  - Koristi niz matrica:  $Q_1, \dots, Q_n \in \mathbb{R}^{m \times m}$  da bi dobio gornjetrougaonu matricu, tj.

$$Q_n \cdots Q_2 Q_1 A = R$$

- Množenje  $i$ -tom matricom anuliraju svi elementi  $i$ -te kolone ispod glavne dijagonale.
- Ovim postupkom QR dekompozicija matrice  $A$  jednaka je:

$$A = Q_1^T Q_2^T \cdots Q_n^T R$$

- Svaka od matrica  $Q_i$  ima formu:

$$Q_i = \begin{pmatrix} I & 0 \\ 0 & F \end{pmatrix}$$

- $I$  kvadrana jedinična matrica dimenzije  $i - 1$ .
- $F$  ortogonalna kvadratna Hausholderova metrica refleksije dimenzije  $m - i + 1$ .
- Ideja konstrukcije matrice  $F$  je da treba da realizuje refleksiju vektora tako da mu svi elementi osim prvog budu 0.

$$Fx = \begin{pmatrix} \|x\|_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- Projekcija na podprostor  $H$ :

$$P_H = I - \frac{vv^T}{v^T v}$$

- Kako transformacija  $F$  treba da pomeri  $x$  dvaput dalje nego projekcija, željena matrica je (za koju se lako pokazuje da je ortogonalna):

$$F = I - 2 \frac{vv^T}{v^T v}$$

- Za vektor  $v$  uzimamo:

$$v = \text{sgn}(x_1) \|x\|_2 e_1 + x$$

Hausholderov algoritam

- for  $i = 1$  to  $n$  do
  - $x = a_{i:m,i}$
  - $v_i = \text{sgn}(x_1) \|x\|_2 e_1 + x$
  - $v_i = v_i / \|v_i\|_2$
  - $a_{i:m,i:n} = a_{i:m,i:n} - 2v_i(v_i^T a_{i:m,i:n})$

Algoritam za množenje vektor  $x$  matricom  $Q$  kada je poznat niz vektora  $v_1, \dots, v_n$ .

- for  $i = n$  to 1 do
  - $x_{i:m} = x_{i:m} - 2 v_i (v_i^T x_{i:m})$

### Singularna dekompozicija

- Matrice blizu singularnim su loše uslovljene, pa SVD pruža smisleno baratanje ovakvim matricama.
- Neka je  $A$  matrica dimenzije  $m \times n$ , pri čemu je  $m \geq n$ .
- Postoji ortogonalna matrica  $U$  dimenzije  $m \times m$ , dijagonalna matrica  $\Sigma$  dimenzije  $m \times n$  i ortogonalna matrica  $V$  dimenzije  $n \times n$ , takva da važi:

$$A = U \Sigma V^T$$

- Matrica  $\Sigma$  ima  $m - n$  vrsta koje se sastoje isključivo od nula pa se mogu zanemariti, uz to uklanjamo i  $m - n$  poslednjih kolona matrice  $U$ , pa dobijamo:

$$A = U' \Sigma' V^T$$

- Dijagonalni elementi matrice  $\Sigma$ ,  $\sigma_i$  za  $i = 1, 2, \dots, n$ , nazivaju se singularnim vrednostima matrice  $A$ .
- Važe sledeća svojstva singularne dekompozicije:
  - Singularne vrednosti matrice  $A$  su sve nenegativne i predstavljaju kvadratne korene sopstvenih vrednosti matrice  $AA^T$  i  $A^T A$ .
  - Rang matrice  $A$  jednak je broju singularnih vrednosti koje nisu jednake nuli.
  - Kolone matrice  $U$  su sopstveni vektori matrice  $AA^T$  i nazivaju se *levim singularnim vektorima* matrice  $A$ .
  - Kolone matrice  $V$  predstavljaju sopstvene vektore matrice  $A^T A$  i nazivaju se *desnim singularnim vektorima* matrice  $A$ .
  - Kolone matrice  $V$  koje odgovaraju nultim singularnim vrednostima predstavljaju ortonormiranu bazu jezgra matrice  $A$ , odnosno prostora svih vektora  $x$ , takvih da važi  $Ax = 0$ .
  - Kolone matrice  $V$  koje odgovaraju nenultim singularnim vrednostima predstavljaju ortonormiranu bazu prostora vrsta matrice  $A$ .
  - Kolone matrice  $U$  koje odgovaraju nenultim singularnim vrednostima predstavljaju ortonormiranu bazu prostora kolona matrice  $A$ .
  - Kolone matrice  $U$  koje odgovaraju nenultim singularnim vrednostima predstavljaju ortonormiranu bazu levog jezgra matrice  $A$ , odnosno prostora svih vektora  $x$ , takvih da važi  $x^T A = 0$ .
  - $\|A\|_2 = \sigma_1$
  - $\|A\|_F = \sqrt{\sum_{i=1}^n \sigma_i^2}$
  - $Cond(A) = \frac{\sigma_1}{\sigma_n}$
- Za računanje singularne dekompozicije potrebno je  $\Theta(mn^2)$  operacija, sa nepovoljnim konstantnim faktorom. Izračunavanje singularne dekompozicije je numerički vrlo stabilno.
- Pomoću singularne dekompozicije može se izvesti ortogonalizacija skupa vektora: Odradimo singularnu dekompoziciju nad matricom čije kolone čine dati vektori, dobijamo ortonormiranu bazu koju čine kolone matrice  $U$ .
- Ako su  $U_i$  kolone matrice  $U$ , a  $V_i$  kolone matrice  $V$ , onda:

$$A = \sum_{i=1}^n \sigma_i U_i V_i^T$$

- Inverz kvadratne matrice pomoću singularne dekompozicije (svaka od singularnih vrednosti  $\sigma_i$ , zamenjena vrednošću  $\frac{1}{\sigma_i}$ ):

$$A^{-1} = V \Sigma^{-1} U^T$$

- U slučaju da  $\Sigma^{-1}$  nije inverz matrice  $\Sigma$ , to znači da je sama matrica  $A$  loše uslovljena.

- Rešavanje sistema  $Ax = b$  (neka je  $x' = V\Sigma^{-1}U^Tb$ ):

$$Ax' = AV\Sigma^{-1}U^Tb = U\Sigma V^TV\Sigma^{-1}U^Tb = U\Sigma\Sigma^{-1}U^Tb$$

- Proizvod  $\Sigma\Sigma^{-1}$  ima nule van dijagonale, jedinice na dijagonali na pozicijama singularnih vrednosti koje nisu nula i nule na ostalim pozicijama.
- $Ax'$  će biti jednak vektoru  $b$  na pozicijama nultih singularnih vrednosti.

## Sopstveni vektori matrica

- Kvadratna matrica  $A$  ima *sopstveni vektor* i sopstvenu vrednost  $\lambda$ , ukoliko važi:

$$Ax = \lambda x \quad x \neq 0$$

- Matrica  $A$  ne menja pravac sopstvenog vektora.
- Različitim sopstvenim vrednostima odgovaraju linearno nezavisni vektori.
- Sopstvene vrednosti jednake su nulama karakterističnog polinoma

$$\det(A - \lambda I).$$

- Svako sopstvenoj vrednosti  $\lambda_i$  odgovara sopstveni potprostor  $V_i$ .
  - \* Neka je  $D$  dijagonalna matrica sopstvenih vrednosti u kojoj se svaka  $\lambda_i$  ponavlja  $\dim(V_i)$  puta, i  $X$  matrica čije su kolone odgovarajući sopstveni vektori, tada važi

$$AX = XD \quad A = XDX^{-1} \quad D = X^{-1}AX$$

- \* Ako je matrica  $V_i$  simetrična, uvek je moguće konstruisati dijagonalnu matricu  $D$ .
- Potpune metode pronalaze sve sopstvene vektore date matrice.
- Delimične metode pronalaze neke sopstvene vektore date matrice.

## Potpune metode

- Pronalazimo nule karakterističnog polinoma  $\det(A - \lambda I)$  i potom rešavamo sistem jednačina  $(A - \lambda I)x = 0$ .
- Jakobijeva metoda zasnovana na primeni matrica rotacije kako bi se anulirali vandijagonalni elementi.
- QR metoda se zasniva na opažanju da:

$$A_1 = A$$

$$A_{i+1} = R_i Q_i$$

- Zbog ortogonalnosti matrica važi:

$$A_{i+1} = R_i Q_i = Q_i^T Q_i R_i Q_i = Q_i^T A_i Q_i$$

- Pod određenim uslovima niz matrica  $R_i$  konvergira matrici koja na dijagonali ima sopstvene vrednosti, a matrica sopstvenih vektor  $X$  se dobija kao proizvod  $Q_1 Q_2 \dots$ .



## Delimične metode

- Delimične metode omogućavaju traženje samo nekih sopstvenih vektora, po pravili dominantnog sopstvenog vektora - onog koji odgovara po modulu najvećoj sopstvenoj vrednosti.
- *Metoda stepenovanja* (power method).
  - Neka su  $x_1, \dots, x_n$  sopstveni vektori matrice  $A \in \mathbb{R}^{n \times n}$ . Neka je  $v_0$  proizvoljan vektor. Sopstveni vektori čine bazu prostora pa važi

$$v_0 = \alpha_1 x_1 + \dots + \alpha_n x_n$$

- Neka je  $v_k = A^k v_0$ , onda

$$v_k = \lambda_1^k \alpha_1 x_1 + \dots + \lambda_n^k \alpha_n x_n$$

- Ako važi  $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$  i  $\alpha_1 \neq 0$ , onda važi:

$$\lambda_2^k \alpha_2 x_2 + \dots + \lambda_n^k \alpha_n x_n = o(\lambda_1^k \alpha_1 x_1)$$

- \* Vektor  $v_k$  konvergira ka sopstvenom vektoru kolinearnom sa  $x_1$ .
- Odgovarajuća sopstvena vrednost je aproksimirana izrazom

$$\frac{v_k^T A v_k}{v_k^T v_k}$$

- Algoritam PageRank se zasniva na metodi stepenovanja.
- Kako izračunati ostale sopstvene vektore?
  - Neka je  $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$  i neka su  $y_1, \dots, y_n$  sopstveni vektori matrice  $A^T$ , skalirani tako da važi  $x_i \cdot y_j = \delta_{ij}$ . Neka je

$$A_1 = A$$

$$A_{i+1} = A_i - \lambda_i x_i y_i^T \quad i = 2, \dots, n$$

- Sopstveni vektori i vrednosti se dobijaju primenom metoda stepenovanja na matrice  $A_i$  i  $A_i^T$ , za  $i = 1, \dots, n$ .
- Ova metoda se naziva *metodom iscrpljivanja*.

## Analiza glavnih komponenti

- Podaci su predstavljeni u visokodimenzionom prostoru.
- Često glavne informacije leže u niskodimenzionim podprostorima.
- Redukcija dimenzionalnosti, radi skladištenja podataka i bržeg izračunavanja mašinskih algoritama.
- Često postoje korelacije i linearne zavisnosti među podacima, pa je eliminisanje redundantnosti jedan o bitnih, a i teških zadatak.
- Racionalan pristup je naći skup vektora duž kojih je varijacija podataka najmanja.
- Kako bi konstruisani vektori bili nekorelisani, oni moraju biti ortogonalni, oni se nazivaju *glavne komponente*.

- Konstruišemo glavne komponente redom, tako da prvi vektor opisuje najvišu varijansu u podacima.
- Prva glavna komponenta se dobija kao rešenje problema

$$\max_{\|v\|=1} \sum_{i=1}^N (x_i \cdot v)^2 = \max_{\|v\|=1} \|Xv\|_2^2 = \max_v \frac{\|Xv\|_2^2}{\|v\|_2^2} = \max_v \frac{\|Xv\|_2}{\|v\|_2}$$

- Prema definiciji matičnih  $p$  normi, maksimalna vrednost veličine koja se maksimizuje je  $\|X\|_2$ , što je takođe najveća singularna vrednost  $\sigma_1$  matrice  $X$ , što je koren po modulu najveće sopstvene vrednosti  $\lambda_1$  matrice  $X^T X$ .

$$\frac{\|Xv\|_2}{\|v\|_2} = \left( \frac{v^T X^T X v}{v^T v} \right)^{1/2} = (\lambda_1 \frac{v^T v}{v^T v})^{1/2} = \sqrt{\lambda_1}$$

- Maksimalna vrednost se dostiže kada je  $v$  dominantan sopstveni vektor matrice  $X^T X$ .
- Sledeću glavnu komponentu dobijamo kada postupak primenimo na  $X_1^T X$ , itd. . .
- Neka su  $v_1, v_2, \dots, v_n$  glavne komponente, svaki podatak  $x$  se može predstaviti pomoću koordinata  $(x \cdot v_1, \dots, x \cdot v_n)$ , odnosno matrica  $X$  se može zameniti matricom  $XV$ .
- Svakoј komponenti se može pridružiti udeo varijanse koju opisuje:

$$\frac{\sigma_i^2}{\sum_{j=1}^n \sigma_j^2}$$

- Udeo varijanse dobijen pomoću prvih  $k$  komponenti je

$$\sum_{i=1}^k \frac{\sigma_i^2}{\sum_{j=1}^n \sigma_j^2}$$

- $X^T X$  predstavlja matricu kovarijanse kolona polazne matrice podataka.
- Podatke možemo standardizovati, tada  $X^T X$ , predstavlja matricu korelacije kolona polazne matrice podataka.
- Da li izvršiti standardizaciju ili ne?
  - Ako su merene na istoj skali, i ne razlikuju se za red veličine standardizacijom gubimo informacije.
  - Inače, treba raditi standardizaciju.

## Retki sistemi linearnih jednačina

- *Retke matrice* se karakterišu malim brojem nenula elemenata.
- Omogućava efikasno skladištenje matrice, kao i efikasno izračunavanje.
- Trivijalan primer rekte matrice je dijagonalna matrice.
  - Za njeno čuvanje je potrebno  $\Theta(n)$ , a rešavanje sistema je  $\Theta(n)$  operacija.

- Slično važi i za matrice oblika:

$$\begin{pmatrix} b_1 & c_1 & 0 & 0 & \dots & 0 & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & \dots & 0 & 0 & 0 \\ 0 & a_3 & b_3 & c_3 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & b_{n-2} & c_{n-2} & 0 \\ 0 & 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & a_n & b_n \end{pmatrix}$$

- Blok-dijagonalne matrice:

$$\begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_m \end{pmatrix}$$

- Sistem  $Ax = b$  rešavamo tako što rešimo pojedinačne podsisteme po blokovima:

$$A_1 x_1 = b_1$$

$$A_2 x_2 = b_2$$

$$\vdots$$

$$A_m x_m = b_m$$

- Složenost rešavanja ovakvog sistema je:  $m\Theta((n/m)^3) = \Theta(n^3)/m^2$ .

- Ostale retke matrice:

- *Trakasta, blok-trougaona, blok-trodijagonalna, blok-dijagonalna sa ivicom, blok-dijagonalna sa dve ivice, blok trougaona sa ivicom, trakasto-trougaona sa ivicom, trakasta sa ivicom, trakasta sa dve ivice...*

## Inkrementalni pristup rešavanju problema linearne algebre

- Da li je moguće, u slučaju malih promene u matrici, efikasno ažurirati dekompoziciju ili inverz?
- *Šerman-Morison-Vudburijeva formula* u slučaju modifikacija niskog ranga:

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I + V^T A^{-1}U)^{-1}V^T A^{-1}$$

- Pri rešavanju sistema  $(A + UV^T)x = b$  važi

$$\begin{aligned} x &= (A + UV^T)^{-1}b \\ &= (A^{-1} - A^{-1}U(I + V^T A^{-1}U)^{-1}V^T A^{-1})b \\ &= A^{-1}b - A^{-1}U(I + V^T A^{-1}U)^{-1}V^T A^{-1}b \\ &= y - ZH^{-1}V^T y \end{aligned}$$

- Pri tome  $y = A^{-1}b$ ,  $Z = A^{-1}U$ ,  $H = I + V^T A^{-1}U$ ,  $y = A^{-1}b$
- $y$  je rešenje sistema  $Ay = b$ , koje može da se izračuna metodom dekompozicije za retke matrice.
- $Z$  može da se izračuna kao rešenje  $p$  retkih sistema

$$Az_1 = U_1$$

$$Az_2 = U_2$$

$$\vdots$$

$$Az_p = U_p$$

- Vektor  $H^{-1}V^T y$  može da se izračuna reštavanjem sistema:

$$Hw = V^T y$$

- Inkrementalno gradimo sistem dodavanjem podataka u primeni metode najmanjih kvadrata.
- Pored primene za lako ažuriranje inverza pomoću prethodnih formula moguće je izvesti i efikasno ažuriranje dekompozicija matrica.
  - Sistem za preporučivanje je jedan idealan primer.

## Matematička optimizacija

Opšti *problem optimizacije* je oblika:

$$\min_{x \in \mathcal{D}} f(x)$$

$$\text{t.d. } g_i(x) \leq 0 \quad i = 1, \dots, M$$

gde je  $f$  *funkcija cilja*, skup  $\mathcal{D}$  *domen*, i  $M$  *funkcija ograničenja*  $g_i$ . Objekat iz domena  $x \in \mathcal{D}$  se naziva *dopustivo rešenje*. Potrebno je među svim dopustivim rešenjima naći ono za koje je vrednost ciljne funkcije najmanja.

- Pronalaženje maksimuma funkcije  $f$  se može svesti na pronalaženje minimuma funkcije  $-f$ .
- Ograničenje  $g(x) = 0$  se može predstaviti pomoću dva ograničenja:  $g(x) \leq 0$  i  $-g(x) \leq 0$ .
- Problemi: raspoređivanje, transport, komunikacija, problemi mašinskog učenja, metode automatskog dizajna hardvera, računarske vizije, robotika, odlučivanja, ekonomije i finansije, biologije, građevine, goe nauka, arheologije, ...
- Podela metoda za rešavanje problema optimizacije po osobinama problema:
  - **Lokalnost**
    - \* Lokalni ili globalni minimumi?
    - \* Lokalne optimizacije su obično egzaktne
    - \* Globalne optimizacije nemaju egzaktne metode, već se rešavaju heuristikama

- **Neprekidnost**
  - \* U zavisnosti od toga da li je domen diskretan ili neprekidan skup?
  - \* Kod diskretnih optimizacija se javlja kombinatorna eksplozija pa se optimizacione metode zasnivaju često na heuristikama.
  - \* Neprekidne optimizacije je obično lako rešiti matematičkom analizom, i obično su te metode efikasne.
- **Diferencijabilnost**
  - \* Da li su funkcija cilja i ograničenja diferencijabilni?
  - \* Ako su neprekidne, koriste gradijent, a ako su još i glatke koriste hesijan kao dodatne informacije o pronalaženju minimuma.
- **Konveksnost**
  - \* Da li su funkcija cilja i ograničenja konveksni?
  - \* Tada imamo jedinstveni optimum, pa se pronalaženje globalnog optimuma svodi na pronalaženje lokalnog optimuma.
- **Prisustvo ograničenja**
  - \* Ako nemamo ograničenja, probleme je moguće rešiti dosta jednostavnijim metodama.

## Primeri praktičnih problema neprekidne matematičke optimizacije

### Neprekidna optimizacija

- U ovom delu se govori o neprekidnoj optimizaciji, tj. domen je neprekidan skup.

### Uslovi optimalnosti

- Pretpostavimo da su funkcije cilja i ograničenja diferencijabilne.
- *Gradijent* funkcije  $f$  u tački  $x$ :

$$\nabla f(x) = \left( \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right)$$

- Gradijent opisuje pravac u kojem funkcija najbrže raste u toj tački.
- U nekoj tački  $x^*$  optimuma, gradijent je nula, tj.

$$\nabla f(x^*) = 0$$

- Tada je tangentna površ je horizontalna.
- Ako je za tačku  $x^*$  važi  $\nabla f(x^*) = 0$ , onda ona ne mora biti tačka optimuma, i takve tačke se nazivaju *stacionarnim*.
- Dva puta diferencijalne funkcije imaju svoj *hesijan*:

$$\nabla^2 f(x) = \left[ \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right]$$

- Da bi stacionarna tačka zaista bila optimum, hesijan u datoj tački mora biti pozitivno ili negativno definitna matrica, tj. mora da važi:

$$h^T \nabla^2 f(x^*) h > 0, \text{ gde } h \neq 0$$

- U slučaju optimizacionih problema sa ograničenjima postoje uslovi optimalnosti *KKT uslovi*:
  - Neka je dat problem

$$\min_{x \in \mathbb{R}^n} f(x)$$

$$\text{t.d. } g_i(x) \leq 0 \quad i = 1, \dots, M$$

$$h_j(x) = 0 \quad j = 1, \dots, L$$

- Neka je  $x^*$  optimalno rešenje i neka su sve funkcije diferencijabilne u  $x^*$ . Ako važe *uslovi regularnosti*, postoje konstante  $\mu_i^*$  i  $\lambda_j^*$  takve da važi:

$$g_i(x^*) \leq 0$$

$$h_j(x^*) = 0$$

$$-\nabla f(x^*) = \sum_{i=1}^M \mu_i^* \nabla g_i(x^*) + \sum_{j=1}^L \lambda_j^* \nabla h_j(x^*)$$

$$\mu_i^* \geq 0$$

$$\mu_i^* g_i(x^*) = 0$$

- Jedan od gornjih uslova možemo posmatrati kao da je  $(x^*, \mu^*, \lambda^*)$  stacionarna tačka *lagranžijana*:

$$L(x, \mu, \lambda) = f(x^*) + \sum_{i=1}^M \mu_i g_i(x) + \sum_{j=1}^L \lambda_j h_j(x)$$

- Uslova regularnosti:
  - Sva ograničenja su afine funkcije
  - Gradijenti aktivnih ograničenja i jednakosnih ograničenja u tački  $x^*$  su linearno nezavisni
  - Sve funkcije u problemu su konveksne i postoji tačka  $x$  takva da je  $h_j(x) = 0$  za sve  $j$  i  $g_i(x) < 0$  za sve  $i$ .
- Dovoljan uslov optimalnosti se može sada definisati preko hesijana: Za svako  $h$  koje zadovoljava  $h^T \nabla g_i(x) = 0$  za svako nejednakosno ograničenje  $g_i$  treba da važi:

$$h^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) h > 0$$

## Metode lokalne optimizacije prvog reda bez ograničenja

- Metode optimizacije prvog reda podrazumevaju sve metode koje kao jedine informacije o funkciji koriste njene vrednosti i vrednosti njenog gradijenta u proizvoljnim tačkama.
- Neka je  $X \subset \mathbb{R}^n$ . Funkcija  $f : X \mapsto \mathbb{R}$  je *Lipšic neprekidna*, ukoliko postoji konstanta  $L$ , takvda da za sve  $x, y \in X$  važi

$$|f(x) - f(y)| \leq L\|x - y\|$$

- Diferencijabilna funkcija  $f : X \mapsto \mathbb{R}$  je *konveksna*, ako za svako  $x, y \in X$  važi:

$$f(x) \geq f(y) + \nabla f(y)^T(x - y)$$

- Funkcija  $f$  je *konkavna* ukoliko je funkcija  $-f$  konveksna.
- Funkcija  $f$  je *jako konveksna*, ukoliko postoji  $m > 0$  i za svako  $x, y \in X$  važi:

$$f(x) \geq f(y) + \nabla f(y)^T(x - y) + \frac{m}{2}\|x - y\|^2$$

- Neformalno, jako konveksna funkcija je konveksna bar koliko i kvadratna funkcija.

- Svojstva konveksnih funkcija:
  - Ako su  $f_1, \dots, f_m$  konveksne funkcije i važi  $w_1 \geq 0, \dots, w_m \geq 0$ , onda je i sledeća funkcija konveksna:

$$w_1 f_1(x) + \dots + w_m f_m(x)$$

- Ako je  $f$  konveksna funkcija,  $A$  matrica i  $b$  vektori odgovarajućih dimenzija, onda je i  $f(Ax + b)$  konveksna funkcija.
- Ako su  $f_1, \dots, f_m$  konveksne funkcije, onda je i sledeća funkcija konveksna:

$$\max\{f_1(x), \dots, f_m(x)\}$$

- \* Isto važi i za supremum nad beskonačnim skupom konveksnih funkcija.
- Kompozicija  $f \circ g$  je konveksna funkcija ako je funkcija  $f$  konveksna i neopadajuća po svim argumentima, a funkcija  $g$  konveksna ili ako je funkcija  $f$  konveksna i nerastuća po svim argumentima, a  $g$  konkavna.
- *Gradijentni spust* je metoda optimizacije prvog reda za diferencijabilna funkcije.
  - Počinjemo od nasumične tačke  $x_0$
  - Svako sledeću tačku računamo na osnovu prethodne:  $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$
  - Kako izabrati parametre  $\alpha_k$ ?
    - \* Konstantne vrednosti:  $\alpha_k = \alpha$ , za svako  $k$ .
    - \* Izbor mora da zadovoljava Robins-Monroove uslove:

$$\sum_{i=0}^{\infty} \alpha_k = \infty \quad \sum_{i=1}^{\infty} \alpha_k^2 < \infty$$

- \* Jedan izbor može biti  $\alpha_k = \frac{1}{k}$ .
- Kriterijum zaustavljanja:
  - \* Određeni broj iteracija;
  - \*  $\|x_{k+1} - x_k\| < \varepsilon$ ;
  - \*  $|f(x_{k+1}) - f(x_k)| < \varepsilon$ ;
  - \*  $\frac{|f(x_{k+1}) - f(x_k)|}{|f(x_0)|} < \varepsilon$ ;
- Za konveksne funkcije sa Lipšic neprekidnim gradijentom, pod Robins-Monroovim uslovima greška  $\|x_k - x^*\|$ , gde je  $x^*$  tačka minimuma, je reda  $O(\frac{1}{k})$ . Ovo implicira da metod konvergira.
- Za jako konveksne funkcije sa Lipšic neprekidnim gradijentom, greška je reda  $O(c^k)$  za neko  $0 < c < 1$ .
- Ostale neprekidne funkcije koje nisu konveksne, gradijentni spust konvergira, ali navedene brzine konvergencija ne važe.
- Na izduženim konturama gradijentni spust pravi zig-zag putanju, kako gradijent ne mora biti pravac najbržeg kretanja ka minimumu.
- Prednost metode gradijentnog spusta su njena jednostavnost i široki uslovi promenljivosti.
- Mane su spora konvergencija, to što je izabran pravac samo lokalno optimalan.
- *Stohastički gradijentni spust* je modifikacija gradijentnog spusta tako što se umesto gradijenta koristi neki slučajni vektor čije je očekivanje kolinearno sa gradijentom i istog je smera.
  - Ima smisla koristiti je kada se funkcija koja se optimizuje može predstaviti kao presek drugih funkcije:

$$f(x) = \frac{1}{n} \sum_{i=1}^N f_i(x)$$

- Korak se onda računa, za nasumično izabevano  $i$ , kao:

$$x_{k+1} = x_k - \alpha_k \nabla f_i(x_k)$$

- Novo  $i$  može da se bira:  $i = (k \bmod N) + 1$
- Još jedan pristup da se za novo rešenje  $x_{k+1}$  uključuje presek nekog podskupa funkcije  $f_i$  (minibatch).
- Za konveksne funkcije sa Lipšic neprekidnim gradijentom greška je  $O(\frac{1}{\sqrt{k}})$
- Za jako konveksne funkcije sa Lipšic neprekidnim gradijentom greška je reda  $O(\frac{1}{k})$ .
- Nekada se gradijent skup za izračunavanje, pa se kod stohastičkog gradijentnog spusta on jeftino aproksimira.
- *Metod inercije* se zasniva na ideji akumuliranje prethodnog gradijenta, pri čemu je značaj starijih gradijenata manji, a novijih veći:

$$d_0 = 0$$

$$d_{k+1} = \beta_k d_k + \alpha_k \nabla f(x_k)$$

$$x_{k+1} = x_k - d_{k+1}$$



- *Nestorovljev ubrzani gradijentni spust* je modifikacija metoda inercije, koja predstavlja asimptotski optimalan algoritam prvog reda za konveksne funkcije:

$$d_0 = 0$$

$$d_{k+1} = \beta_k d_k + \alpha_k \nabla f(x_k - \beta_k d_k)$$

$$x_{k+1} = x_k - d_{k+1}$$

- Za konveksne funkcije sa Lipšic neprekidnim gradijentom, greška je reda  $O(\frac{1}{k^2})$ .

### Metode lokalne optimizacije drugog reda bez ograničenja

- Metode optimizacije drugog reda pored vrednosti funkcija i gradijenta, koriste hesijan.
- Kako gradijent pruža informaciju o brzini promene funkcije duž različitih koordinatnih pravaca, tako hesijan pruža informaciju o brzini promene gradijenta duž različitih koordinatnih pravaca.
- *Njutnov metod*:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

$$x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k)$$

- Za jako konveksne funkcije sa Lipšic neprekidnim hesijanom, greška je reda  $O(c^{2^k})$ , za neko  $0 < c < 1$ , što je neuporedivo brže od metoda prvog reda.
- Neka je funkcija koja se minimizuje kvadratna:

$$f(x) = \frac{1}{2} x^T A x + b^T x + c$$

- Odgovarajući gradijent:  $\nabla f(x) = b + Ax$
- Odgovarajući hesijana:  $\nabla^2 f(x) = A$

- Korak Njutnove metode:

$$x_1 = x_0 - A^{-1}(b + Ax) = -A^{-1}b$$

$$\nabla f(x_1) = \nabla f(-A^{-1}b) = b + A(-A^{-1}b) = 0$$

- Iz prethodnog razmatranja imamo da je gradijent nula, pa je smo dobili stacionarnu tačku.
- Ako je  $f$  konveksna funkcija, tj. matrica  $A$  je pozitivno semidefinitna, sigurno se radi o minimumu.
- Ako je  $f$  konkavna funkcija, tj. matrica  $A$  je negativno semidefinitna, sigurno se radi o maksimumu.
- U svim ostalim slučajevima radi se o sedlenim tačkama.
- Njutnova metoda traži nulu gradijenta, a ne minimum funkcije.
  - Kasnijim razmatranjem to će biti ili minimum ili maksimum ili sedlena tačka.

- Njutnovom metodom se vrši niz uzastopnih minimizacija lokalnih kvadratnih aproksimacija funkcije.
- Prednost Njutnove metode je brza konvergencija
- Mana je memorijski zahtevno skladištenje hesijana, i zahtev za strogu konveksnost funkcije.
- *Kvazi-Njutnove metode* se zasnivaju na aproksimaciju inverza hesijana na osnovu gradijenata.
- *BFGS* (Brojden-Flečer-Goldfarb-Šano):

$$x_{k+1} = x_k - H_k^{-1} \nabla f(x_k)$$

- $H_k^{-1}$  aproksimirana simetrična matrica.
- BFGS pretpostavlja i slaganje gradijenata funkcije  $f$  i njene kvadratne aproksimacije  $\bar{f}_k$ :

$$\bar{f}_k(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T H_k (x - x_k)$$

$$\nabla \bar{f}_k(x) = \nabla f(x_k) + H_k (x - x_k)$$

- Gradijenti se slažu u tački  $x_k$ .

$$\nabla f(x_k) + H_k (x_{k-1} - x_k) = \nabla f(x_{k-1})$$

- Ovaj uslov se oslanja na matricu  $H_k$ , što je nepoželjno, pošto je aproksimirana  $H_k^{-1}$ , pa se uslov transformiše u ekvivalentan:

$$H_k^{-1} (\nabla f(x_k) - \nabla f(x_{k-1})) = x_k - x_{k-1}$$

- Dodatno se zahteva da  $H_k^{-1}$  predstavlja rešenje narednog optimizacionog problema:

$$\min_{H^{-1}} \|H^{-1} - H_{k-1}^{-1}\|_2^2$$

$$\text{t.d. } H^{-1} (\nabla f(x_k) - \nabla f(x_{k-1})) = x_k - x_{k-1}$$

$$H^{-1T} = H^{-1}$$

- Ispostavlja se da ovaj problem ima rešenje u zatvorenoj formi, koje se brzo izračunava.
- BFGS ima red greške između  $O(c^k)$  i  $O(c^{2^k})$ . Može se očekivati da ova metoda bude sportija od Njutnove, ali brža od metoda prvog reda.
- Ne rešava problem memorije, to radi LBFGS (low memory BFGS).

### Linijaska pretraga

- Linijaska pretraga se zasniva na izboru dužine koraka. Pretražuje odabranu duž pravca za najboljom ili maka povoljnom dužinom koraka.
- Pretpostavimo da je izbor pravca spusta, tj. da važi  $\nabla f(x)^T d < 0$ , gde je  $d$  pravac.

- Egzaktna linijska pretraga u tački  $x_k$ :

$$\min_{\alpha \geq 0} f(x_k + \alpha d)$$

- Da li se ovaj problem može rešiti analitički ili ne?
- U praksi se retko koristi.
- Kako izabrati odgovarajuću vrednosti  $\alpha$ ?
  - Neka je  $\alpha_k = \alpha_0 \beta^k$  za  $k > 0$ ,  $\alpha_0 > 0$  i  $\beta \in (0, 1)$ . Linijska pretraga se bira najmanje  $k$ , odnosno najveće  $\alpha_k$  za koje važe Armihov uslov:

$$f(x + \alpha_k d) \leq f(x) + \alpha_k \nabla f(x)^T d.$$

- Da je ovaj postupak izbora vrednosti  $\alpha$  završava u konačnom vremenu?
  - Za dovoljno malo  $\alpha$  važi:

$$f(x + \alpha d) \approx f(x) + \alpha \nabla f(x)^T d$$

- Kako  $\alpha_k$  eksponencijalno opada, za dovoljno veliko  $k$  važi  $\alpha_k < \alpha$ .
- Kako je  $d$  pravac spusta, važi:

$$f(x) + \alpha \nabla f(x)^T d < f(x) + \alpha^* \nabla f(x)^T d$$

- Za dovoljno malo  $\alpha$  važi:

$$f(x + \alpha f) < f(x) + \alpha^* \nabla f(x)^T d$$

### Metode lokalne optimizacije sa ograničenjima

- Uz prisustvo ograničenja minimum funkcije ne mora biti jednak pravom minimumu funkcije.
- Takođe, ukoliko ne postoji minimum funkcije bez ograničenja, on može postojati ukoliko su ograničenja pristna.
- Najjednostavnije klasa problema sa ograničenjima su linearni problemi, odnosno problemi *linearnog programiranja*.
- Najpoznatiji metod rešavanja problema linearnog programiranja je *simpleks algoritam*.
  - Ima eksponencijalnu složenost, ali u praksi je često efikasan.
  - Postoji i algoritmi sa polinomijalnom složenošću.
- U slučaju konveksnog skupa dopustivih rešenja, moguće je primeniti mehanizam *projektovanog gradijenta*.

$$\min_{u \in U} \|x - u\|_2 = P_U(x)$$

$$x_{k+1} = P_U(x_k - \alpha_k \nabla f(x_k))$$

- Kako rešiti problem projektovanja?
  - Optimizacijom? Ne pokazuje se toliko efikasno
  - Metodi zasnovana na *kaznenim funkcijama* (algoritam *logaritamska barijera*)

- Opšti problem minimizacije rešavamo iterativno tako što se u  $k$ -toj iteraciji rešava problem (za početnu tačku uzima se rešenje prethodne iteracije):

$$\min_x f(x) + \frac{1}{\mu_k} \sum_{i=1}^L -\log(-g_i(x))$$

- Kada je  $g_i(x)$  blisko nuli, vrednost kaznene funkcije  $-\log(-g_i(x))$  je veliki pozitivan broj.
- Povećavanjem parametra  $\mu$  se omogućava smanjenje uticaja kaznene funkcije.

## Diskretna optimizacija

- Diskretna optimizacija podrazumeva diskretnost nekog od elemenata optimizacionog problema (domena, funkcije cilja).
- Metode se posmatraju kao algoritmi pretrage na prostoru potencijalnih rešenja.
- Egzaktna pretraga garantuje pronalaženje optimuma.
- Heuristička pretraga ne pruža nikakvu garanciju.

## Egzaktne metode

- Problemi koje rešavaju egzaktne metode su vrlo često  $NP$ -teški.
- Zbog toga ove metode imaju eksponencijalne ili veće vremenske složenosti.
- *Grananje i ograničavanje* (Branch and bound)
  - Prostor rešenja se može deliti na dva ili više delova koji u uniji čine ceo prostor.
  - Iscrpna pretraga, ali ako je moguće uštedu u vremenu treba izvršiti odsecanjem podstabla pretrage.
  - Zasniva se na brzom određivanju donjih granica vrednosti funkcije cilja: kada je donja granica nekog od potprostora veća od najniže vrednosti pronađene u toku pretrage, celo podstablo koje odgovara tom potprostoru se može zanemariti.

Algoritam:

- Nekom heuristikom odrediti početno dopustivo rešenje  $x$ ,
- Neka je  $B = f(x)$ ,  $s = x$  i  $Q = [P]$ .
- Ponavljati dok  $Q \neq \emptyset$ 
  - Uzeti instancu  $I$  iz reda  $Q$
  - Ukoliko važi  $single(I)$  i ako je  $x$  jedino rešenje instance  $I$  i važi  $f(x) < B$ , onda dodeliti  $B = f(x)$  i  $s = x$  i preskočiti ostatak iteracije.
  - Neka je  $[I_1, \dots, I_n] = branch(I)$
  - Svaku instancu  $I_j$  za koju važi  $bound(I_i) < B$  staviti u red
- Vratiti ( $s, B$ )

## Heurističke metode

- Heurističke metode ne garantuju optimalnost.
- *Metaheuristike* su šabloni po kojima se kreiraju *heuristike* za konkretan problem.
- *Populacione* metaheuristike se zasnivaju na održavanju populacije dopustivih rešenja koja se paralelno menjaju, popravljaju, kobinuju, interaguju i slično. (genetski algoritmi, kolonija mrava, roj čestica, ...)
- Druge predstavljaju održanje jednog dopusivog rešenja. (simulirano kaljenje, tabu pretraga, metod promenljivih okolina).
- Mehanizam intenzifikacije popravljaju tekuće rešenje ili tekuća rešenje.
- Mehanizam diverzifikacije pokušava da se izvuče iz lokalno minimuma.
- *Metod promenljivih okolina* (VNS — variable neighbourhood search)
  - Održava jedno rešenje
  - Koristi metod pronalaženja lokalnog optimuma (metod intenzifikacije)
  - Koristi *razmrđavanje* (shaking) (metod diverzifikacije)
- $(N)_i(x)$ , za  $i = 1, \dots, K$  za svako dopustivo rešenje  $x \in D$ .

Redukovana metoda promenljivih okolina:

- Inicijalizovati dopustivo rešenje  $x$
- Ponavljati naredne korake sve dok nije ispunjen kriterijum zaustavljanja:
  - Neka je  $k = 1$ .
  - Ponavljati dok važi  $k < K$ 
    - \* Razmrđavanje: nasumice generisati tačku  $x' \in \mathcal{N}_k(x)$
    - \* Kretanje: ukoliko je  $f(x') < f(x)$ , neka je  $x = x'$  i  $k = 1$ , a u suprotnom, neka je  $k = k + 1$

Metoda promenljivih okolina:

- Inicijalizovati dopustivo rešenje  $x$
- Ponavljati naredne korake sve dok nije ispunjen kriterijum zaustavljanja:
  - Neka je  $k = 1$ .
  - Ponavljati dok važi  $k < K$ 
    - \* Razmrđavanje: nasumice generisati tačku  $x' \in \mathcal{N}_k(x)$
    - \* Lokalna pretraga: primeniti neki metod lokalne optimizacije počevši od  $x'$  i označiti rezultat sa  $x''$
    - \* Kretanje: ukoliko je  $f(x'') < f(x)$ , neka je  $x = x''$  i  $k = 1$ , a u suprotnom, neka je  $k = k + 1$

Opšta metoda promenljivih okolina se dobija kada se za mehanizam lokalne pretrage uzme *spust sa promenljivih okolinama* (variable neighbourhood descent):

- Neka je  $l = 1$
- Ponavljati dok važi  $l < L$ 
  - Pretraga: naći najbolje  $x' \in \mathcal{N}_l(x)$
  - Kretanje: ukoliko je  $f(x') < f(x)$ , neka je  $x = x'$  i  $l = 1$ , a u suprotnom, neka je  $l = l + 1$
- Okoline  $\mathcal{N}$  i  $N$  se ne moraju podudarati.