Studentska natjecanja 2024.

Numerička matematika na natjecanjima

11. 4. 2024.
Matko Ljulj
mljulj[at]math.hr

Uvod

Neki od vas nisu slušali Numeričku matematiku u 4. semestru, neki od vas slušaju sada, neki od vas je već jesu slušali. Neovisno o emocijama koje gajite prema tom kolegiju, postoje zadatci na natjecanjima koji, uz ne jako široko znanje teorije numeričke matematike, olakšavaju rješavanje zadataka. Ti zadatci često su na natjecanjima okarakterizirani teškima. Vjerojatno zato što naš fakultet ima natprosječan pristup numeričkoj matematici, pa su naše "osnove" drugima "za one koji žele znati više".

1 Interpolacijski polinom

Znamo da je skup svih polinoma stupnja manjeg ili jednakog n (uz standardne operacije zbrajanja i množenja skalarom) vektorski prostor $\mathcal{P}_n$ dimenzije $n+1$. Stoga je za očekivati da možemo naći polinom upravo takvog stupnja koji zadovoljava $p(x_i) = y_i$, $i = 0, \dots, n$, za neke različite realne $x_0, \dots, x_n$ i realne $y_0, \dots, y_n$ jer smo zadali točno $n+1$ uvjeta na taj polinom p. Taj problem zovemo *problem interpolacije*. Takav polinom uistinu postoji.

Jedan način za dokazati da postoji takav (jedinstveni) polinom je korištenjem činjenice da je skup $\{1, x, x^2, \dots, x^n\}$ baza za $\mathcal{P}_n$. Koeficijente $a_0, \dots, a_n$ u polinomu

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

nalazimo rješavanjem sustava

$$\begin{bmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Sustav ima rješenje jer je matrica sustava regularna za različite x_i-jeve, njena determinanta je tzv. Vandermondeova determinanta.

Drugi način da vidimo da postoji interpolacijski je da odaberemo bolju bazu. Za vektorski prostor $\mathcal{P}$ postoji mnogo odabira baza, pa možemo odabrati onu bazu polinoma $\{l_0, l_1, \dots, l_n\}$ za koju vrijedi $l_i(x_j) = \delta_{i,j}$. Ta baza dana je s

$$l_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}.$$

Baza nema lijep zapis, no nepoznati skalari koji množe elemente baze u problemu interpolacije su trivijalni: naime polinom

$$p(x) = \sum_{i=0}^{n} l_i(x) y_i$$

zaista zadovoljava $p(x_i) = y_i$ za sve $i \in \{0, 1, \dots, n\}$. Zadnja formula formula je za Lagrangeov interpolacijski polinom.

Napomena 1. *Iako ime sugerira drugačije, Lagrangeov (pa kasnije i Newtonov) interpolacijski polinom nisu polinom koji se razlikuje od interpolacijskog polinoma danog u standardnoj bazi. To je uvijek jedan jedinstveni interpolacijski polinom, dan u drugačijoj bazi. Zato bi ispravnije bilo reći Lagrangeov zapis interpolacijskog polinoma.*

Dakle, sad smo obradili dva slučaja: rješenje interpolacijskog problema možemo tražiti u kanonskoj bazi ili Lagrangeovoj bazi. S jedne strane imamo vrlo jednostavnu bazu, ali kompliciran način traženja skalara, dok s druge strane imamo kompliciranu bazu, polinome s kojima je teško baratati u zadatcima, ali s druge strane koeficijenti su vrlo lijepi. Zato ćemo ponuditi i treću mogućnost, koji bi trebao biti kompromis ove dvije baze.

Ponovno fiksirajmo različite $x_0, \dots, x_n$. Biramo bazu

$$\{1, (x - x_0), (x - x_0)(x - x_1), \dots, (x - x_0)\dots(x - x_{n-1})\},$$

te je formula za interpolacijski polinom dana s

$$p(x) = [y_0] + [y_0, y_1](x - x_0) + [y_0, y_1, y_2](x - x_0)(x - x_1) + \dots + [y_0, y_1, \dots, y_n](x - x_0)\dots(x - x_{n-1}).$$

Izrazi oblika $[y_i, y_{i+1}, \dots, y_{j-1}, y_j]$ zovu se *podijeljene razlike*, te se nalaze rekurzivno po broju članova:

$$[y_i] := y_i, \ [y_i, \dots, y_j] := \frac{[y_{i+1}, \dots, y_j] - [y_i, \dots, y_{j-1}]}{x_j - x_i}.$$

Ovakav polinom nazivamo *Newtonovim interpolacijskim polinomom.* Primjer izgradnje podijeljenih razlika vidjet ćemo u primjeru.

Primjer 1.1 . *Nađimo Newtonov interpolacijski polinom koji prolazi točkama* $(1, 2)$, $(3, 4)$, $(8, -1)$, $(10, 0)$. *Baza polinoma je* $\{1, (x - 1), (x - 1)(x - 3), (x - 1)(x - 3)(x - 8)\}$. *Koeficijente nalazimo iz tablice i algoritmu nalik na punjenju Pascalovog trokuta. Popunjavamo razinu po razinu prema formuli, samo što je konvencija da se trokut postavlja s bazom položenom vertikalno, te popunjavamo slijeva nadesno. Prvo pišemo* x_i*-jeve, desno u odgovarajućim retcima njihove vrijednosti (*y_i*-jeve), a onda radimo trokut.*

$$\begin{array}{llllll}
x_0 = 1 & y_0 = 2 & & & & \\
 & & [y_0, y_1] = \frac{4-2}{3-1} = 1 & & & \\
x_0 = 3 & y_1 = 4 & & [y_0, y_1, y_2] = \frac{-1-1}{8-1} = \frac{-2}{7} & & \\
 & & [y_1, y_2] = \frac{-1-4}{8-3} = -1 & & [y_0, y_1, y_2, y_3] = \frac{\frac{3}{14} - \frac{-2}{7}}{10-1} = \frac{-1}{18} & \\
x_2 = 8 & y_2 = -1 & & [y_1, y_2, y_3] = \frac{\frac{1}{2} - (-1)}{10-3} = \frac{3}{14} & & \\
 & & [y_2, y_3] = \frac{0-(-1)}{10-8} = \frac{1}{2} & & & \\
x_3 = 10 & y_3 = 0 & & & &
\end{array}$$

Rješenje čitamo iz najgornjeg reda:

$$p(x) = 2 + 1(x - 1) - \frac{2}{7}(x - 1)(x - 3) - \frac{1}{18}(x - 1)(x - 3)(x - 8).$$

Dakle, prednost Newtonovog algoritma je što nudi neki kompromis između komplicirane formule za bazu polinoma i formula za koeficijente. S druge strane, postoji još jedna prednost: pomoću tog algoritma lako možemo zadati vrijednosti polinoma u točkama x_i, ali i vrijednosti nekih derivacija proizvoljnog reda. Općenito vrijedi sljedeći teorem:

Teorem 1 (Hermite)**.** *Neka su* $x_0, \dots, x_k$ *različiti realni brojevi,* $r_0, \dots, r_n$ *neki nenegativni cijeli brojevi, te* $y_i^{(j)}$ *ukupno* $n := r_0 + \dots + r_k + k + 1$ *realnih brojeva, gdje indeksi idu* $0 \leqslant j \leqslant r_i$, $0 \leqslant i \leqslant n$. *Tada postoji jedinstveni polinom* p *stupnja manjeg ili jednakog* $n - 1$ *takav da je* $p^{(j)}(x_i) = y_j^{(j)}$, *za sve* $0 \leqslant j \leqslant r_i$, $0 \leqslant i \leqslant n$.

Algoritam za pronalaženje takvog polinoma analogan je onom za traženje Newtonovog interpolacijskog polinoma, uz napomenu da svaku točku x_i u tablicu stavljamo $r_i + 1$ puta, te podijeljena razlika u kojoj se m puta pojavljuje isti broj $[y_i, \dots, y_i]$ mijenja se s $\dfrac{f^{(m-1)}(x_i)}{(m-1)!}$.

Primjer 1.2 . *Nađimo polinom najnižeg stupnja za koji vrijedi* $f(0) = 0$, $f'(0) = 1$, $f''(0) = -2$, $f(1) = 2$. *Baza u kojoj tražimo rješenje je* $\{1, (x - 0), (x - 0)^2, (x - 0)^3\}$, *a podijeljene razlike tražimo rekurzivno*

$$\begin{array}{lllll}
0 & [y_0] = 0 & & & \\
 & & [y_0, y_1] = f'(0) = 1 & & \\
0 & [y_0] = 0 & & [y_0, y_1, y_2] = \frac{f''(0)}{2} = -1 & \\
 & & [y_1, y_2] = f'(0) = 1 & & [y_0, y_1, y_2, y_3] = \frac{1-(-1)}{1-0} = 2 \\
0 & [y_0] = 0 & & [y_1, y_2, y_3] = \frac{2-1}{1-0} = 1 & \\
 & & [y_2, y_3] = \frac{2-0}{1-0} = 2 & & \\
1 & [y_1] = 2 & & &
\end{array}$$

Rješenje ponovno čitamo iz prvog reda:

$$f(x) = 0 + x + (-1)x^2 + 2x^3.$$

Zadatak 1.
Dokaži da za sve normirane polinome $f(x)$ stupnja $n-1$ i različite realne brojeve $a_1, \cdots, a_n$ vrijedi

$$\sum_{i=1}^{n} \frac{f(a_i)}{\prod_{j \neq i}(a_i - a_j)} = 1.$$

Rješenje.
Izraz u zadatku neodoljivo podsjeća na Lagrangeov interpolacijski polinom, i on se sam poziva da riješi zadatak. Sjetimo se, interpolacijski polinom p stupnja manjeg ili jednakog $n-1$ za funkciju f u točkama $a_1, \ldots, a_n$ je

$$p(x) = \sum_{i=1}^{n} \frac{\prod_{j \neq i}(x - a_i)}{\prod_{j \neq i}(a_i - a_j)} f(a_i).$$

U našem slučaju, funkcija f koju želimo interpolirati već jest polinom stupnja najviše $n-1$, pa je jednaka svom interpolacijskom polinomu. Jedina informacija koju još nismo iskoristili je da je normiran, a to rješava zadatak. Naime, lako se vidi da ako želimo u Lagrangeovom polinomu naći vodeći koeficijent, treba primijetiti da brojnik svake bazne funkcije ima vodeći koeficijent 1 i to uvrstiti u sumu.

Također, Lagrangeov interpolacijski polinom ponekad je korisno pisati i u formi

$$p(x) = \sum_{k=1}^{n} \frac{\omega(x)}{(x - a_k)\omega'(a_k)} p(a_k),$$

gdje je $\omega(x) = (x - a_1) \cdots (x - a_n)$. Upravo to ćemo iskoristiti u sljedećem zadatku:

Zadatak 2.
Neka su $x_1, \ldots, x_n$ različiti realni brojevi i koji nisu jednaki niti 0 niti -1. Ako je $\omega(x) = (x - x_1) \cdots (x - x_n)$, dokaži

$$\sum_{k=1}^{n} \frac{x_k^n \omega(1/x_k)}{\omega'(x_k)(1 + x_k)} = (-1)^{n-1}(1 - x_1 \cdot x_n).$$

Rješenje.
Ovaj zadatak je manje direktan od prošlog. Za početak, nije nam dan niti jedan polinom koji bismo trebali interpolirati. Ono što nam hintira kako odrediti taj polinom su brojnici u pribrojnicima sume na lijevoj strani. Dakle, naivni pokušaj je da stavimo $p(x) = x^n \omega(1/x)$. No, to je polinom stupnja n, za koji interpolacijska formula nije egzaktna:

$$x^n \omega(1/x) = (1 - x x_1) \cdots (1 - x x_n).$$

Popraviti ideju možemo na dva načina: ili povećati broj interpolacijskih točaka (recimo ubaciti 0 ili -1), ili smanjiti stupanj gore napisanom polinomu tako da mu oduzmemo neki polinom stupnja n s istim vodećim koeficijentom. Napravit ćemo ovo drugo: definiramo

$$p(x) := x^n \omega(1/x) + (-1)^n (1 - x_1 \cdots x_n) r(x),$$

gdje je $r(x)$ neki normirani polinom stupnja n koji ćemo kasnije definirati.
Uvrštavanjem toga u Lagrangeov interpolacijski polinom, i evaluacijom u točki -1 (što je logično gledajući nazivnik), dobivamo jednakost

$$(-1)^n \frac{p(-1)}{\omega(-1)} = \sum_{k=1}^{n} \frac{x_k^n \omega(1/x_k)}{\omega'(x_k)(1 + x_k)} + \sum_{k=1}^{n} \frac{(-1)^{n-1} x_1 \cdots x_n r(x_k)}{\omega'(x_k)(1 + x_k)}.$$

Ovo vrijedi za svaki normirani polinom $r(x)$ stupnja n. Da bismo dobili formulu iz zadatka, treba odabrati najpovoljniji. To možemo ili tako da namjestimo da je $p(-1) = 0$, ili bolje, tako da stavimo da su svi $r(x_k) = 0$. Dakle, stavimo $r(x) = \omega(x)$. Sada je cijela druga suma na desnoj strani jednaka nuli, a uvrštavajući u definiciju polinoma $p(x)$ vrijednost -1 dobivamo traženu jednakost.

2 Čebiševljevi polinomi

Sljedeća tema predavanja su najpoznatije familije rekurzivno zadanih polinoma.

Definicija 1. *Čebiševljevi polinomi prve vrste zadani su rekurzivnom relacijom*

$$T_0(x) = 1,\ T_1(x) = x,\ T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), n \in \mathbb{N},$$

dok su Čebiševljevi polinomi druge vrste zadani sa

$$U_0(x) = 1,\ U_1(x) = 2x,\ U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), n \in \mathbb{N}.$$

Popis nekoliko prvih Čebiševljevih polinoma:

$T_0(x) = 1$	$U_0(x) = 1$
$T_1(x) = x$	$U_1(x) = 2x$
$T_2(x) = 2x^2 - 1$	$U_2(x) = 4x^2 - 1$
$T_3(x) = 4x^3 - 3x$	$U_3(x) = 8x^3 - 4x$
$T_4(x) = 8x^4 - 8x^2 + 1$	$U_4(x) = 16x^4 - 12x^2 + 1$

Najbitnija njihova svojstva proizlaze iz sljedeće leme.

Lema 1. *Za sve $\varphi \in \mathbb{R}$ vrijedi*

$$T_n(\cos\varphi) = \cos n\varphi,\ U_{n-1}(\cos\varphi) = \frac{\sin n\varphi}{\sin\varphi}.$$

Dokaz. Uvedimo pokratu $y := \cos\varphi$. Dokaz se temelji na indukciji i jednostavnim trigonometrijskim identitetima. Primjerice, za T_n imamo (za $U_n(x)$ ide slično)

$$\cos(n+1)\varphi + \cos(n-1)\varphi = 2\cos n\varphi \cdot \cos\varphi,$$

odakle je

$$\cos(n+1)\varphi = 2\cos n\varphi\cos\varphi - \cos(n-1)\varphi = 2yT_n(y) - T_{n-1}(y) = T_{n+1}(y).$$

Iz te leme dobivamo zatvorene formule za nultočke:

$$x_k(T_n) = \cos\left(\frac{2k-1}{2n}\pi\right), \qquad x_k(U_n) = \cos\left(\frac{k}{n+1}\pi\right)\ k = 1, \dots, n.$$

Također, vrlo lako se vidi i da su točke u kojima jedni Čebiševljevi polinomi postižu lokalne ekstreme upravo nultočke onog drugog polinoma (s odgovarajućim pomakom na indeksima).

Dodatno, za Čebiševljeve polinome prve vrste vrijedi i sljedeće:

$$t \neq 0: \quad T_n\left(\frac{1}{2}\left(t + \frac{1}{t}\right)\right) = \frac{1}{2}\left(t^n + \frac{1}{t^n}\right).$$

Drugo svojstvo koje se također koristi na natjecanjima:

Teorem 2. *Za sve normirane polinome stupnja najviše n vrijedi nejednakost*

$$\max_{x\in[-1,1]} |f(x)| \geqslant \frac{1}{2^{n-1}},$$

gdje se jednakost postiže samo za $T_n(x)/2^{n-1}$.

Dokaz. Pretpostavimo suprotno, postoji polinom $w(x)$ stupnja najviše n za koji vrijedi suprotna nejednakost. Definirajmo $f(x) = T_n(x)/2^{n-1} - w(x)$. Kako je $|w(x)| < 1/2^{n-1}$, u svim točkama u kojima polinom $T_n(x)/2^{n-1}$ poprima apsolutnu vrijednost vrijednost, funkcija $f_n(x)$ ima jasno definiran predznak u ekstremima od $T_n(x)$. Konkretno:

$$f(x) > 0 \text{ za } x = \cos\frac{2k\pi}{n},\ f(x) < 0 \text{ za } \ x = \cos\frac{(2k+1)\pi}{n}.$$

Tako smo za polinom f koji je stupnja najviše $n-1$ našli $n+1$ točku u kojoj on promijeni predznak, pa po neprekidnosti ima barem n nultočaka. Kontradikcija.

Generalizacija gornje tvrdnje:

Teorem 3 (Markov brother's inequality)**.** *Neka je $p(x)$ polinom s realnim koeficijentima stupnja n. Tada je*

$$\max_{-1\leqslant x\leqslant 1} |p^{(k)}(x)| \leqslant \frac{n^2(n^2-1)(n^2-4)\cdots(n^2-(k-1)^2)}{1\cdot 3\cdot 5\cdots(2k-1)} \max_{-1\leqslant x\leqslant 1} |p(x)|.$$

Jednakost se postiže za polinom $T_n(x)$.

3 SVD dekompozicija

Na predmetima iz linearne algebre naučili smo da se neke matrice daju dijagonalizirati. Dijagonalu joj čine svojstvene vrijednosti. To svojstvo je istinito za dosta malu skupinu matrica: matrica mora biti kvadratna, i to ne bilo kakva, a tek ako je normalna znamo da joj svojstveni vektori mogu biti izabrani tako čine ortonormiranu bazu.

Teorem 4. *Dana je matrica $\mathbf{A}\in\mathbb{R}^{m\times n}$. Tada postoje unitarne matrice $\mathbf{U}\in\mathbb{R}^{m\times m}, \mathbf{V}\in\mathbb{R}^{n\times n}$ te "dijagonalna" $\mathbf{\Sigma}\in\mathbb{R}^{m\times n}$ tj. jednaka nuli osim na vrijednostima $\Sigma_{i,i}$ gdje ima vrijednosti $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_{\min\{m,n\}}$ takva da je*

$$\mathbf{A}=\mathbf{U\Sigma V}^T.$$

Gornji rastav zovemo singularna dekompozicija matrice (eng. singular value decomposition), ili kraće SVD. Vrijednosti $\sigma_1,\ldots,\sigma_{\min\{m,n\}}$ nazivamo singularne vrijednosti matrice $\mathbf{A}$, a stupce matrica $\mathbf{U}$ i $\mathbf{V}$ zovemo lijevi i desni singularni vektori pridruženi tim vrijednostima.

Napomena 2. *Pogledajmo neka svojstva na prvu:*

- *Dakle, ova dekompozicija postoji za sve matrice $\mathbf{A}$. Dapače, isto se može napraviti i za matrice iz $\mathbb{C}^{m\times n}$. Tada se u jednakosti transponiranje mijenja s hermitskim adjungiranjem, no matrica Σ i dalje ima realne (i nerastuće) vrijednosti. Matrica ne mora biti kvadratna, ne mora imati spektralnu dekompoziciju i slično.*
- *Kako svaka $\mathbf{A}$ dopušta singularnu dekompoziciju, možemo pisati primjerice*

  $$\mathbf{AA}^T=\mathbf{U\Sigma V}^T(\mathbf{U\Sigma V}^T)^T=\mathbf{U}(\mathbf{\Sigma\Sigma}^T)\mathbf{U}^T.$$

  Matrica $\mathbf{\Sigma\Sigma}^T$ je dijagonalna, elementi su joj kvadrati singularnih vrijednosti. Primijetimo da smo time dokazali da su singularne vrijednosti kvadrati svojstvenih vrijednosti matrice $\mathbf{AA}^T$. Posebno, singularne vrijednosti normalne matrice su apsolutne vrijednosti njenih svojstvenih vrijednosti.

Pogledajmo još neka svojstva SVD-a.

Propozicija 1. *Neka je $\mathbf{A}=\mathbf{U\Sigma V}^T$ singularna dekompozicija matrice $\mathbf{A}$. Tada je rang matrice $\mathbf{A}$ jednak broju ne-nul singularnih vrijednosti. Slika od $\mathbf{A}$ razapeta je stupcima matice $\mathbf{U}$ koji ne pripadaju singularnim vrijednostima nula, a jezgra stupcima matice $\mathbf{U}$ koji pripadaju singularnim vrijednostima nula.*

Prisjetimo se matričnih normi. Za broj $1\leqslant p<\infty$ i vektor $\mathbf{x}\in\mathbb{R}^n$ definiramo normu reda p kao

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p},$$

a za $p=\infty$

$$\|\mathbf{x}\|_\infty := \max_{i=1,\ldots,n} |x_i|.$$

Općenitu matricu $\mathbf{A}\in\mathbb{R}^{m\times n}$ isto možemo promatrati kao vektor i na jednak način definirati normu. Iako time dobivamo matrične norme (i sve norme na konačnodimenzionalnom prostoru su ekvivalentne), to se rijetko radi u praksi. Jedina iznimka je za $p=2$, koju zovemo Frobeniusova norma. Ona se veže i uz odgovarajući Frobeniusov skalarni produkt

$$\langle\mathbf{A},\mathbf{B}\rangle=\mathrm{tr}(\mathbf{A}^T\mathbf{B}).$$

Umjesto toga, kada pričamo o matričnoj normi reda p, mislimo na sljedeću definiciju:

$$\|\mathbf{A}\|_p=\max_{\mathbf{x}\neq 0}\frac{\|\mathbf{Ax}\|_p}{\|\mathbf{x}\|_p}=\max_{\|\mathbf{x}\|=1}\|\mathbf{Ax}\|_p=\max_{\|\mathbf{x}\|\leqslant 1}\|\mathbf{Ax}\|_p.$$

Ovakve norme zovemo operatorskima, i općenito za njih vrijedi

$$\|\mathbf{AB}\| \leqslant \|\mathbf{A}\| \cdot \|\mathbf{B}\|$$

za sve pravokutne matrice za koje je dobro definiran umnožak $\mathbf{AB}$.

Što se tiče operatorskih normi, najčešće se koriste za $p = 1, 2, \infty$. I dok za $p = 1, \infty$ postoje jednostavne formule za račun odgovarajuće norme (maksimalna suma apsolutnih vrijednosti u nekom stupcu/retku), za $p = 2$ nije tako jednostavno za računati. Zato je korisno da postoji sljedeći rezultat.

Propozicija 2. *Neka je $\mathbf{A} = \mathbf{U\Sigma V}^T$ singularna dekompozicija matrice $\mathbf{A}$. Tada je Frobenisova norma matrice $\mathbf{A}$ jednaka korijenu sume kvadrata singularnih vrijednosti matrice $\mathbf{A}$. Matrična norma reda 2 ($\|\mathbf{A}\|_2$) jednaka je singularnoj vrijednosti σ_1.*

Primijetimo još jednu posljedicu SVD-a, koja se da nevjerojatno lako dokazati.

Primjer 3.3 . *Dokažimo da se svaka matrica ranka r da zapisati kao zbroj r matrica ranga 1.*
Ako označimo s $\mathbf{u}_1, \ldots, \mathbf{u}_r$ i $\mathbf{v}_1, \ldots, \mathbf{v}_r$ prvih r stupaca matrica $\mathbf{U}$ i $\mathbf{V}$ iz SVD-a od $\mathbf{A}$, te uočimo da je matrica $\mathbf{xy}^T$ matrica ranga 1, tada je lako za vidjeti da vrijedi

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T.$$

Postoje još mnoga zanimljiva svojstva ove dekompozicije. Jedno od njih korisno je i u "stvarnom" životu.

Teorem 5 (Ekhard, Young, Mirsky). *Dana je matrica $\mathbf{A} \in \mathbb{M}^{m\times n}$ ranga r te prirodan broj $k < r$. Tada za svaku matricu $\mathbf{K} \in \mathbb{M}^{m\times n}$ ranga k vrijedi*

$$\|\mathbf{A} - \mathbf{K}\|_2 \geqslant \sigma_{k+1},$$

gdje je σ_i i-ta po veličina singularna vrijednost matrice $\mathbf{A}$. Jednakost se postiže za matricu $\mathbf{K} = \mathbf{U\Sigma}_k\mathbf{V}^T$, gdje su $\mathbf{U}$, $\mathbf{\Sigma}$ i $\mathbf{V}$ matrice iz SVD za matricu $\mathbf{A}$, a Σ_k dobivena brisanjem svih osim prvih k singularnih vrijednosti iz $\mathbf{\Sigma}$.

Zamislimo da se umjesto o matrici radi o nekoj slici pospremljenoj na računalu. Ona je ogromna matrica (reda N) u kojima su upisane primjerice RGB vrijednosti u odgovarajućem pikselu. Jedan način komprimiranja njene slike je da umjesto te slike (= matrice $\mathbf{A}$) pamtimo neku njenu aproksimaciju $\mathbf{A}_k$, uz $k << N$, tj trebamo samo zapamtiti prvih k singularnih vrijednosti i prvih k singularnih vektora, što je mnogo manje podataka. S druge strane, kvaliteta slike će biti nešto manja, no ovakav algoritam daje začuđujuće dobre rezultate. Analogan algoritam može se napraviti i za probleme predikcije.

Sva ova priča o SVD-u dana je zato što je na izbornom natjecanju za Vojtech Jarnik jednom dan zadatak koji se s SVD-om da lako riješiti.

Zadatak 3.
Dana je fiksna realna kvadratna matrica $\mathbf{A}$ reda n. Pretpostavimo da ortogonalna matrica $\mathbf{Q}$ reda n maksimizira $\mathrm{tr}(\mathbf{QA})$. Dokažite da je tada matrica $\mathbf{QA}$ simetrična.

Rješenje.
Neka je $\mathbf{A} = \mathbf{U\Sigma V}^T$ SVD dekompozicija matrice $\mathbf{A}$. Ako je $\mathbf{A}$ nul-matrica, tvrdnja je trivijalna pa pretpostavimo da je $k \geqslant 1$ najveći indeks takav da je $\sigma_k > 0$. Iz poznatog svojstva traga $\mathrm{tr}(\mathbf{MN}) = \mathrm{tr}(\mathbf{NM})$ i uz supstituciju $\mathbf{R} = \mathbf{V}^T\mathbf{QU}$ imamo

$$\mathrm{tr}(\mathbf{QA}) = \mathrm{tr}(\mathbf{QU\Sigma V}^T) = \mathrm{tr}(\mathbf{V}^T\mathbf{QU\Sigma}) = \mathrm{tr}(\mathbf{R\Sigma}).$$

Primijetimo da je $\mathbf{R} = [r_{i,j}]$ zapravo ortogonalna matrica reda n koja maksimizira

$$\mathrm{tr}(\mathbf{R\Sigma}) = \sum_{i=1}^{n} \sigma_i r_{i,i} = \sum_{i=1}^{k} \sigma_i r_{i,i}.$$

Budući da je $r_{i,i} \leqslant 1$, gornji izraz je najviše $\sum_{i=1}^{k} \sigma_i$ i jednakost se postiže samo u slučajevima kada je $r_{i,i} = 1$ za $i = 1, 2, \ldots, k$, što ima za posljedicu da je prvih k stupaca od $\mathbf{R}$ upravo prvih k vektora kanonske baze od $\mathbb{R}^n$. Odavde se odmah vidi $\mathbf{R\Sigma} = \mathbf{\Sigma}$ te je zato

$$\mathbf{QA} = \mathbf{QU\Sigma V}^T = \mathbf{VR\Sigma V}^T = \mathbf{V\Sigma V}^T,$$

što je očito simetrična matrica.

SVD je općenito moćna dekompozicija, no nije svemoguća, i daleko od toga da je jedina. Još neke dekompozicije (pored SVD i spektralne) koje ćete čuti tijekom studija su: dekompozicija Choleskog, Schurova dekompozicija, polarna dekompozicija, LU faktorizacija, QR faktorizacija, Jordanova dekompozicija, ...

4 Spektri matrica AB i BA, argument neprekidnosti

Dokažimo sljedeću zanimljivu tvrdnju:

Teorem 6. *Dane su matrice* $\mathbf{A} \in \mathbb{M}^{m\times n}$, $\mathbf{B} \in \mathbb{M}^{n\times m}$. *Tada matrice* $\mathbf{AB}$ *i* $\mathbf{BA}$ *dijele nenul svojstvene vrijednosti računajući kratnost.*

Pokazat ćemo dva dokaza, s time da će nam drugi biti zanimljiviji.

Dokaz. Možemo se uvjeriti da vrijedi sljedeća matrična jednakost:

$$\begin{bmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{AB} & \mathbf{0} \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B} & \mathbf{BA} \end{bmatrix}$$

Odavde zaljučujemo da su matrice $\begin{bmatrix} \mathbf{AB} & \mathbf{0} \\ \mathbf{B} & \mathbf{0} \end{bmatrix}$ i $\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{B} & \mathbf{BA} \end{bmatrix}$ slične, pa dijele spektar. Njihovi spektri su spektri matrica $\mathbf{AB}$ i $\mathbf{BA}$, i još neke nule.

Dokaz. Za tren simulirajmo gornji dokaz u jednostavnijem formatu. Opet ćemo dokazati koristeći sličnost da matrice dijele spektar.

Pretpostavimo da je $m = n$, i da je $\mathbf{A}$ regularna. Onda možemo pisati

$$\mathbf{AB} = \mathbf{A}(\mathbf{BA})\mathbf{A}^{-1},$$

odakle zaključujemo sličnost.

Pokušajmo ovo iskoristiti u dokazu i za singularnu $\mathbf{A}$. Ideja je da ako je $\mathbf{A}$ singularna, da je za sve dovoljno male $\varepsilon > 0$ matrica $\mathbf{A}_\varepsilon := \mathbf{A} + \varepsilon\mathbf{I}$ regularna. To je posljedica toga što $\mathbf{A}_\varepsilon$ nije regularna samo u svojstvenim vrijednostima od $\mathbf{A}$, kojih je konačno mnogo. Gornji rezultat možemo iskoristiti tako da zaključimo da za sve male $\varepsilon > 0$ matrice $\mathbf{A}_\varepsilon\mathbf{B}$ i $\mathbf{BA}_\varepsilon$ dijele spektar (dapače, slične su). Želimo nekako pustiti $\varepsilon \to 0$. Najlakše je to koristeći karakteristične polinome: za sve male $\varepsilon > 0$

$$\det(\lambda\mathbf{I}_n - \mathbf{A}_\varepsilon\mathbf{B}) = \det(\lambda\mathbf{I}_n - \mathbf{BA}_\varepsilon).$$

Za fiksan λ, gornje funkcije po ε su neprekidne (jer se radi o polinomima), pa možemo obje strane pustiti u nulu, i zaključiti

$$\det(\lambda\mathbf{I}_n - \mathbf{AB}) = \det(\lambda\mathbf{I}_n - \mathbf{BA}),$$

čime je dokaz za kvadratne matrice gotov. Ako je BSO $m < n$, onda nadodamo odgovarajuće nule, tj. iskoristimo dokazano za $\mathbf{A}_1 = \begin{bmatrix} \mathbf{A} \\ \mathbf{0} \end{bmatrix}$, $\mathbf{B}_1 = \begin{bmatrix} \mathbf{B} & \mathbf{0} \end{bmatrix}$.

Ovaj način dokazivanja možemo koristiti i inače. Za općenitu tvrdnju

- prvo dokažemo za regularne matrice;
- zamijenimo singulanu $\mathbf{A}$ s $\mathbf{A}_\varepsilon := \mathbf{A} + \varepsilon\mathbf{I}$;
- naći način kako zaključiti tvrdnju puštanjem $\varepsilon \to 0$.

Pogledajmo to na još jednom primjeru:

Zadatak 4.
Dokažite da za adjunkte kvadratnih matrica $\mathbf{A}$ i $\mathbf{B}$ vrijedi $\operatorname{adj}(\mathbf{AB}) = \operatorname{adj}(\mathbf{B})\operatorname{adj}(\mathbf{A})$.

Rješenje.
Možda ovo ima i direktan dokaz, koristeći definiciju adjunkte, ali vjerojatno je odvratan. Koristimo formulu za inverz (regularne) matrice $\mathbf{X}$ vrijedi formula $\mathbf{X}^{-1} = (1/\det\mathbf{X}) \cdot \operatorname{adj}(\mathbf{X})$.

Ako želimo iskoristiti gornju formulu, prvo trebamo pretpostaviti da su obje matrice $\mathbf{A}$ i $\mathbf{B}$ regularne. U tom slučaju im je i umnožak regularan te imamo

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1} \implies \frac{1}{\det(\mathbf{AB})}\operatorname{adj}(\mathbf{AB}) = \frac{1}{\det(\mathbf{B})}\operatorname{adj}(\mathbf{B})\frac{1}{\det(\mathbf{A})}\operatorname{adj}(\mathbf{A}).$$

Tvrdnja slijedi direktno iz Binet-Cauchyjevog teorema.

Neka su sada $\mathbf{A}$ i $\mathbf{B}$ singularne. Tada su $\mathbf{A}_\varepsilon$ i $\mathbf{B}_\varepsilon$ regularne za sve dovoljno male ε.

Iz definicije $\operatorname{adj}(\mathbf{X})$ (na mjestu (i,j) se nalazi $(-1)^{i+j}$ pomnoženo s determinantom $(n-1)\times(n-1)$ matrice $\mathbf{X}$ iz koje smo prekrižili i-ti stupac i j-ti redak) vidimo da je ona polinom svojih matričnih elemenata. Posljedično, $\varepsilon \to \operatorname{adj}(\mathbf{A}_\varepsilon)$ je neprekidna funkcija svojih matričnih elemenata. Zato u jednakosti

$$\operatorname{adj}(\mathbf{A}_\varepsilon\mathbf{B}_\varepsilon) = \operatorname{adj}(\mathbf{B}_\varepsilon)\operatorname{adj}(\mathbf{A}_\varepsilon)$$

isto vrijedi da se na svakom mjestu (i,j) nalazi neprekidna funkcija po ε, pa ju smijemo pustiti u 0, čime dobivamo tvrdnju.

Domaća zadaća

Da bi se zadaća smatrala riješenom, treba riješiti barem četiri od navedenih sedam zadataka. Predajete ju putem maila (slikanu ili skeniranu), rok je 2 tjedna od predavanja.

DZ 1. *Za fiksan $n \geqslant 2$ i realan parametar a odrediti sve različite kompleksne brojeve $x_1, \ldots, x_n$ takve da vrijednost izraza*

$$M(a) = \max_{k=1,\ldots,n} |p_k(a)|, \ p_k(a) := \frac{\omega(a)}{\omega'(a)(x-x_k)}$$

bude najmanja moguća.

DZ 2. *Neka su $a_0, a_1, \ldots, a_n$ realni brojevi takve da je za sve $x \in [-1,1]$ $|a_0 + a_1x + \cdots a_nx^n| \leqslant 1$. Dokaži da je tada i za sve $x \in [-1,1]$ $|a_n + a_{n-1}x + \cdots a_0x^n| \leqslant 2^{n-1}$.*

DZ 3. *Neka su $x_1, \ldots, x_n$, $n \geqslant 2$ različiti realni brojevi iz intervala $[-1,1]$. Dokaži da je*

$$\frac{1}{t_1} + \cdots + \frac{1}{t_1} \geqslant 2^{n-2},$$

gdje je $t_i = \prod_{j\neq i}|x_j - x_i|$.

DZ 4. *Dane pravokutne matrice $\mathbf{A}, \mathbf{B} \in \mathbb{M}^{m\times n}$ za koje vrijedi $\mathbf{A}^T\mathbf{A} = \mathbf{B}^T\mathbf{B}$. Dokaži da postoji ortogonalna $\mathbf{Q} \in \mathbb{M}^{m\times m}$ takva da je $\mathbf{A} = \mathbf{QB}$.*

DZ 5. *Dane su simetrične matrice $\mathbf{A}$ i $\mathbf{B}$. Dokažite da vrijedi $\operatorname{tr}(\mathbf{AB})^2 \leqslant \operatorname{tr}(\mathbf{A}^2\mathbf{B}^2)$. Kada vrijedi jednakost?*

DZ 6. *U ovisnosti o parametrima $x_1, \ldots, x_n, y_1, \ldots, y_n$ izračunajte vrijednost determinante*

$$\begin{vmatrix} 1+x_1y_1 & x_1y_2 & \ldots & x_1y_n \\ x_2y_1 & 1+x_2y_2 & \ldots & x_2y_n \\ \vdots & \vdots & \vdots & \vdots \\ x_ny_1 & x_ny_2 & \ldots & 1+x_ny_n \end{vmatrix}.$$

DZ 7. *Dokažite Ekhard–Young–Mirskyjev teorem.*

DZ 8. *Za proizvoljnu kvadratnu matricu $\mathbf{A}$ reda n kojoj su svojstvene vrijednosti $\lambda_1, \lambda_2, \ldots, \lambda_n$ odredite (uz dokaz) svojstvene vrijednosti matrice $\operatorname{adj}(\mathbf{A})$.*