

## Kapitola 15

# Několik aplikací

### Diskrétní a rychlá Fourierova transformace

Diskrétní Fourierova transformace spočívá ve změně reprezentace polynomu s koeficienty v nějakém tělese  $\mathbf{T}$ . Obvyklá reprezentace polynomu stupně  $n - 1$

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1}$$

sestává ze zadání koeficientů  $a_0, a_1, \dots, a_{n-1} \in \mathbf{T}$ . Stejně tak ale můžeme zadat polynom  $f$  pomocí jeho hodnot v  $n$  různých bodech  $x_0, x_1, \dots, x_{n-1} \in \mathbf{T}$ . Předepíšeme-li hodnoty  $f(x_i) = b_i \in \mathbf{T}$  pro  $i = 0, \dots, n - 1$ , pak podle Tvzení 4.15 existuje právě jeden polynom  $f$  stupně  $n - 1$  s koeficienty v tělese  $\mathbf{T}$ , který má v bodech  $x_0, \dots, x_{n-1}$  předepsané hodnoty. Jeho koeficienty  $a_0, \dots, a_{n-1}$  najdeme jako řešení soustavy lineárních rovnic

$$\begin{aligned} a_0 + a_1x_0 + \cdots + a_{n-1}x_0^{n-1} &= b_0 \\ a_0 + a_1x_1 + \cdots + a_{n-1}x_1^{n-1} &= b_1 \\ &\vdots \\ a_0 + a_1x_{n-1} + \cdots + a_{n-1}x_{n-1}^{n-1} &= b_{n-1}. \end{aligned}$$

Matice této soustavy je Vandermondova matice

$$V_{x_0, x_1, \dots, x_{n-1}} = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} \\ 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^{n-1} \end{pmatrix}.$$

Jednoznačnost polynomu  $f$  vyplývá z regularity Vandermondovy matice v případě, že body  $x_0, x_1, \dots, x_{n-1}$  jsou navzájem různé, kterou jsme dokázali v Úloze 9.3.

Předpokládejme nyní, že máme vynásobit dva polynomy  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$  a  $g(x) = c_0 + c_1x + c_2x^2 + \dots + c_{n-1}x^{n-1}$  stupně nejvýše  $n - 1$ . Jejich součin  $fg$  je polynom stupně nejvýše  $2n - 2$ . Budeme jej znát, pokud budeme znát  $2n - 1$  koeficientů součinu  $fg$  nebo hodnoty součinu  $fg$  ve  $2n - 1$  různých bodech  $x_0, x_1, \dots, x_{2n-2}$ . Známe-li hodnoty obou polynomů  $f(x_i) = b_i$  a  $g(x_i) = c_i$  v těchto  $2n - 2$  bodech, můžeme snadno spočítat hodnoty součinu  $fg(x_i) = b_i c_i$  v těchto bodech. Potřebujeme k tomu pouze  $2n - 1$  násobení, zatímco při výpočtu koeficientů  $d_j$  při klasické reprezentaci součinu

$$(fg)(x) = d_0 + d_1x + \dots + d_{2n-2}x^{2n-2}$$

pomocí vztahů

$$d_k = \sum_{i=0}^k a_i c_{k-i}, \quad k = 0, 1, \dots, 2n - 2,$$

potřebujeme celkem

$$2(1 + 2 + \dots + n) = n(n + 1)$$

násobení a řádově stejný počet sčítání. Násobení polynomů je tak mnohem rychlejší, pokud máme polynomy zadány pomocí jejich hodnot v dostatečně mnoha bodech.

Problém samozřejmě spočívá v ceně přechodu od zadání polynomu pomocí koeficientů k jeho zadání pomocí hodnot v předepsaných bodech a obráceně. Hodnotu polynomu  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$  v bodě  $x_i$  můžeme spočítat pomocí vyjádření

$$\begin{aligned} f(x_i) &= a_0 + a_1x_i + a_2x_i^2 + \dots + a_{n-1}x_i^{n-1} = \\ &= a_0 + x(a_1 + x(a_2 + x(\dots + x(a_{n-1})\dots))). \end{aligned}$$

Tomuto vyjádření se říká *Hornerovo schéma*. Vyžaduje  $n - 1$  násobení a stejný počet sčítání. Potřebujeme-li spočítat hodnotu polynomu  $f$  v  $n$  různých bodech, musíme tak použít celkem  $n(n - 1)$  násobení a  $n(n - 1)$  sčítání. To je řádově stejný počet násobení jako při přímém výpočtu koeficientů součinu a úplně stejný počet sčítání. Protože násobení je mnohem náročnější na čas než sčítání, můžeme říct, že výpočet hodnoty polynomu stupně  $n - 1$

v  $n$  bodech je přibližně stejně časově náročný jako spočítat součin dvou polynomů stupně  $n - 1$  pomocí přímého výpočtu koeficientů.

Všimněme si, že hodnoty polynomu  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$  v bodech  $x_0, \dots, x_{n-1}$  můžeme spočítat jako souřadnice součinu matic

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^{n-1} \\ 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1} & x_{n-1}^2 & \cdots & x_{n-1}^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix},$$

což můžeme rovněž zapsat ve tvaru

$$(f(x_0), f(x_1), \dots, f(x_{n-1}))^T = V_{x_0, x_1, \dots, x_{n-1}}(a_0, a_1, \dots, a_{n-1})^T.$$

Výpočtu hodnot polynomu  $f$  v bodech  $x_0, \dots, x_{n-1}$  budeme říkat *evaluační transformace* a budeme ji označovat  $T_{x_0, x_1, \dots, x_{n-1}}(f)$ .

Pokud naopak chceme z hodnot polynomu  $f$  stupně  $n - 1$  v  $n$  bodech najít koeficienty polynomu, musíme vyřešit soustavu  $n$  lineárních rovnic o  $n$  neznámých s regulární maticí, což při výpočtu pomocí Gaussovy eliminace vyžaduje dokonce řádově  $n^3$  násobení a řádově stejný počet sčítání. K tomu je ještě třeba přidat cenu výpočtu prvků Vandermondovy matice, která je maticí soustavy.

Abychom mohli využít “rychlejšího” násobení polynomů při jejich zadání pomocí hodnot v dostatečně mnoha bodech, potřebujeme mnohem rychlejší algoritmy pro výpočet hodnoty polynomu v bodě než je výpočet pomocí Hornerova schématu. Stejně tak potřebujeme mnohem rychlejší algoritmus pro výpočet koeficientů polynomu, známe-li jeho hodnoty v dostatečně mnoha bodech, než je algoritmu založený na Gaussově eliminaci použité na Vandermondovu matici. Tyto algoritmy si nyní ukážeme.

Podstata urychlení Hornerova schématu spočívá ve vhodném výběru bodů  $x_0, x_1, \dots, x_{n-1} \in \mathbf{T}$ , ve kterých hodnoty polynomu  $f$  počítáme. Pokud předpokládáme, že  $n = 2k$  je sudé číslo, pak můžeme polynom  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$  napsat ve tvaru

$$f(x) = g(x^2) + xh(x^2),$$

kde

$$g(y) = a_0 + a_2y + a_4y^2 + \dots + a_{n-2}y^{k-1}$$

a

$$h(y) = a_1 + a_3y + a_5y^2 + \dots + a_{n-1}y^{k-1}.$$

Oba polynomy  $g, h$  mají poloviční stupeň oproti polynomu  $f$ , počítáme ale jejich hodnoty stále v  $n$  bodech  $x_0^2, \dots, x_{n-1}^2$  a protože polynomy jsou dva, žádné snížení počtu násobení zatím není vidět. Pokud ale navíc předpokládáme, že body  $x_0, \dots, x_{n-1}$  jsou zvolené *symetricky*, tj. že platí

$$x_{k+i} = -x_i$$

pro  $i = 0, 1, \dots, k-1$ , tak vidíme, že jsme nejenom snížili stupeň obou polynomů na polovinu, ale rovněž jsme snížili na polovinu počet bodů, ve kterých musíme počítat hodnoty obou polynomů. To už vede k podstatnému snížení počtu potřebných algebraických operací, jak ukazuje následující tvrzení.

**Tvrzení 15.1** *Předpokládáme, že  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$  je polynom stupně  $n-1$  s koeficienty v tělese  $\mathbf{T}$ , číslo  $n = 2k$  je sudé, a prvky  $x_0, x_1, \dots, x_{n-1} \in \mathbf{T}$  splňují podmínku  $x_{k+i} = -x_i$  pro  $i = 0, 1, \dots, k-1$ . Označíme-li  $C(n)$  počet operací, které jsou třeba k výpočtu hodnot nějakého polynomu v  $n$  bodech  $x_0, x_1, \dots, x_{n-1}$ , pak*

$$C(1) = 0, \quad \text{a} \quad C(n) = 2 \cdot C\left(\frac{n}{2}\right) + 4 \cdot \frac{n}{2}.$$

**Důkaz.** Protože platí  $x_{k+i}^2 = x_i^2$  pro  $i = 0, \dots, k-1$ , potřebujeme spočítat hodnoty polynomů  $g, h$  pouze v  $k = n/2$  bodech. To vyžaduje celkem  $2 \cdot C(n/2)$  operací. Kromě toho potřebujeme  $n/2$  násobení pro výpočet čtverců  $x_i^2$  pro  $i = 0, 1, \dots, n/2-1$ . Dále  $n/2$  násobení,  $n/2$  sčítání a  $n/2$  odčítání, abychom zkombinovali hodnoty  $g(x_i^2)$  a  $h(x_i^2)$  do hodnot  $f(x_i)$  pro  $i = 0, 1, \dots, n-1$ .  $\square$

Vhodná volba evaluačních bodů je popsána v následující definici.

**Definice 15.2** *Prvek  $\omega$  tělesa  $\mathbf{T}$  se nazývá primitivní  $n$ -tá odmocnina z 1, pokud  $\omega^n = 1 (= \omega^0)$  a mocniny*

$$\omega, \omega^2, \dots, \omega^{n-1}$$

*jsou navzájem různé. Množina prvků  $\{1, \omega, \omega^2, \dots, \omega^{n-1}\}$  se nazývá Fourierovy body. Evaluační transformace  $T_{1, \omega, \dots, \omega^{n-1}}$  ve Fourierových bodech  $1 = \omega^0, \omega, \omega^2, \dots, \omega^{n-1}$  se nazývá diskrétní Fourierova transformace (DFT).*

**Příklad 15.3** *1.Prvek*

$$\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$$

je primitivní  $n$ -tá odmocnina z 1 v tělese komplexních čísel  $\mathbf{C}$ , jak vyplývá z Moivreovy věty. Fourierovy body

$$\omega^k = \cos \frac{k \cdot 2\pi}{n} + i \sin k \cdot 2\pi n, \quad k = 0, 1, \dots, n-1$$

tvoří vrcholy pravidelného  $n$ -úhelníka vepsaného do jednotkové kružnice tak, aby bod  $(1, 0)$  byl jedním z vrcholů tohoto  $n$ -úhelníka.

2. V konečném tělese  $\mathbf{Z}_{17}$  je prvek  $\omega = 4$  primitivní čtvrtou odmocninou z 1, neboť v tomto tělese platí  $4^4 = 256 = 1 \pmod{17}$  a mocniny

$$4^1 = 4, \quad 4^2 = 16, \quad 4^3 = 13$$

jsou navzájem různé. Vandermondova matice

$$V_{1,4,16,13} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 16 & 13 \\ 1 & 16 & 1 & 16 \\ 1 & 13 & 16 & 4 \end{pmatrix}$$

popisuje diskrétní Fourierovu transformaci  $T_{1,4,16,13}$ .

3. V tělese  $\mathbf{Z}_{41}$  je číslo 14 primitivní osmou odmocninou z 1, odpovídající Fourierovy body jsou 1, 14, -9, -3, -1, -14, 9, 3.

Je-li  $\omega \in \mathbf{T}$  primitivní  $n$ -tá odmocnina z 1 pro sudé  $n = 2k$ , pak  $n$  Fourierových bodů  $1, \omega, \omega^2, \dots, \omega^{n-1}$  splňuje podmínku

$$(\omega^{k+i})^2 = \omega^{2k+2i} = \omega^{n+2i} = \omega^n \omega^{2i} = 1 \cdot (\omega^i)^2 = (\omega^i)^2.$$

pro  $i = 0, 1, \dots, k-1$ . Navíc je očividně  $\omega^2$  primitivní  $(n/2)$ -tá odmocnina z 1 a v případě, že  $n = 4l$  splňuje  $k = n/2$  sudých mocnin  $1, \omega^2, \omega^4, \dots, \omega^{2k-1}$ , také stejnou podmínku symetrie

$$((\omega^2)^{l+i})^2 = \omega^{4l+4i} = \omega^{4i} = ((\omega^2)^i)^2.$$

K výpočtu hodnot polynomů  $g, h$  v  $n/2$  bodech  $\omega_i^2, i = 0, 1, \dots, k = n/2$  můžeme použít stejný trik s výpočtem hodnot 4 polynomů ve  $l = n/4$  bodech. A tak dále.

Pokud je tedy  $n = 2^m$  a  $\omega \in \mathbf{T}$  primitivní  $n$ -tá odmocnina z 1, můžeme rekurzivně spočítat hodnotu diskrétní Fourierovy transformace  $T_{1,\omega,\dots,\omega^{n-1}}$ . Tomuto výpočtu se říká *rychlá Fourierova transformace (FFT)*. Následující výpočet ukazuje, jak moc FFT zrychluje výpočet hodnoty polynomu v  $n = 2^m$  bodech.

**Tvrzení 15.4** Je-li  $n = 2^m$  a  $\omega \in \mathbf{T}$  primitivní  $n$ -tá odmocnina z 1, pak FFT vyžaduje

$$C(n) = 2 \cdot n \cdot \log_2 n$$

aritmetických operací.

**Důkaz.** Opakovaným použitím Tvrzení 15.1 dostaneme

$$\begin{aligned} C(n) = C(2^m) &= 2 \cdot C(2^{m-1}) + 4 \cdot 2^{m-1} = \\ &= 2(2 \cdot C(2^{m-2}) + 4 \cdot 2^{m-2}) + 4 \cdot 2^{m-1} = \\ &= 2^2 \cdot C(2^{m-2}) + 2 \cdot 4 \cdot 2^{m-1} = \\ &= 2^3 \cdot C(2^{m-3}) + 3 \cdot 4 \cdot 2^{m-1} = \\ &\vdots \\ &= 2^m \cdot C(1) + m \cdot 4 \cdot 2^{m-1} = \\ &= \log_2 n \cdot 4 \cdot \frac{n}{2} = \\ &= 2 \cdot n \cdot \log_2 n. \end{aligned}$$

□

Diskrétní Fourierova transformace a její spojitá verze nazývaná prostě Fourierova transformace jsou hojně používány v elektrotechnice, jedno využití spočívá v odstraňování šumu v signálech. V počítačové algebře je rychlá Fourierova transformace používána v systémech pro symbolickou manipulaci k násobení polynomů velkého stupně.

K tomu potřebujeme znát inverzní diskretní Fourierovu transformaci, pomocí které ze znalosti hodnoty polynomu v  $n = 2^m$  bodech spočítáme jeho koeficienty. Tato inverzní diskretní Fourierova transformace

$$T_{1,\omega,\dots,\omega^{n-1}}^{-1}$$

odpovídá výpočtu inverzní matice  $V_{1,\omega,\dots,\omega^{n-1}}^{-1}$  k matici  $V_{1,\omega,\dots,\omega^{n-1}}$ . Lze ukázat, že

$$V_{1,\omega,\dots,\omega^{n-1}}^{-1} = n^{-1} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \dots & \omega^{-(n-1)} \\ 1 & \omega^{-2} & \omega^{-4} & \dots & \omega^{-2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{-(n-1)} & \omega^{-2(n-1)} & \dots & \omega^{(n-1)^2} \end{pmatrix}.$$

Protože  $\omega^{-1} \in \mathbf{T}$  je rovněž primitivní  $n$ -tá odmocnina z 1, můžeme pro výpočet inverzní diskretní Fourierovy transformace rovněž použít rychlou Fourierovu transformaci.

Rychlou Fourierovu transformaci poprvé uveřejnili pánové J.W. Cooley a J.W. Tuckey v roce 1965 a od té doby se stala jedním z nejpoužívanějších algoritmů.

### Vyhledávač Google

Popularita vyhledávač Google spočívá také v tom, že uspořádá vyhledané stránky k danému dotazu podle “důležitosti”. Ukážeme si, jak je tato “důležitost” jednotlivých webových stránek definována a jak ji lze vypočítat.

Řekneme, že *důležitost nějaké webové stránky je přímo úměrná součtu důležitostí stránek, které na ni odkazují*. To na první pohled vypadá na definici kruhem. Zkusíme ji ale vyjádřit matematicky. Označíme  $x_i$  důležitost  $i$ -té webové stránky pro  $i = 1, 2, \dots, N$ , kde  $N$  je celkový počet webových stránek, které bereme v úvahu. Dále ještě potřebujeme nějak zachytit strukturu vzájemných odkazů mezi jednotlivými stránkami. Pokud  $j$ -tá stránka odkazuje na  $i$ -tou stránku, tak napíšeme, že

$$a_{ij} = 1.$$

V opačném případě, tj. pokud  $j$ -tá stránka neodkazuje na  $i$ -tou stránku, položíme

$$a_{ij} = 0.$$

Struktura vzájemných odkazů mezi jednotlivými webovými stránkami je tak popsána pomocí čtvercové matice  $\mathbf{A} = (a_{ij})$  řádu  $N$ . Tuto matici můžeme nazvat *matice incidence webu*.

Je-li  $K$  konstanta přímé úměrnosti použitá v neformální definici důležitosti  $i$ -té stránky, pak můžeme tuto definici formalizovat následovně:

$$x_i = K \cdot \sum_{j=1}^N a_{ij} x_j \quad \text{pro } i = 1, 2, \dots, N.$$

Přepsáním dostaneme

$$\sum_{j=1}^N a_{ij} x_j = \frac{1}{K} \cdot x_i \quad \text{pro } i = 1, 2, \dots, N.$$

Označíme  $\mathbf{x} = (x_1, x_2, \dots, x_N)^T$  vektor důležitostí jednotlivých stránek, pak poslední rovnost zapíšeme ve tvaru

$$\mathbf{A}\mathbf{x} = \frac{1}{K}\mathbf{x}.$$

Vidíme tedy, že koeficient přímé úměrnosti  $K$  je převrácenou hodnotou nějakého vlastního čísla matice  $\mathbf{A}$  (matice incidence webu) a vektor důležitostí  $\mathbf{x}$  je vlastním vektorem příslušným tomuto vlastnímu číslu.

Matice incidence webu  $\mathbf{A}$  má  $N$  (komplexních) vlastních čísel a ke každému vlastnímu číslu existuje nekonečně mnoho nenulových odpovídajících vlastních vektorů. Potřebujeme proto nějak vybrat vhodné vlastní číslo matice  $\mathbf{A}$ , nejlépe tak, aby bylo kladné reálné, a k němu najít vhodný vlastní vektor, opět by bylo nejlepší, aby to byl vektor s kladnými reálnými souřadnicemi, nebo aspoň s nezápornými.

V tomto místě nastupuje *Perronova-Frobeniova teorie nezáporných matic*. Jako první si uvedeme definici spektrálního poloměru matice.

**Definice 15.5** *Je-li  $\mathbf{B}$  čtvercová komplexní matice, pak číslo*

$$\rho(\mathbf{A}) = \max\{|\lambda| : \lambda \in \sigma(\mathbf{B})\}$$

*nazýváme spektrální poloměr matice  $\mathbf{B}$ . Kružnici se středem v počátku a poloměrem  $\rho(\mathbf{B})$  nazýváme spektrální kružnice matice  $\mathbf{B}$ .*

Spektrální poloměr každé čtvercové komplexní matice je tedy nezáporné reálné číslo. Takovým maticím říkáme *kladná (pozitivní) matice*. Pokud jsou všechny prvky matice nezáporná reálná čísla, pak mluvíme o *nezáporné matici*. Matice, jejichž prvky jsou kladná reálná čísla, mají následující důležitou vlastnost.

**Věta 15.6** *Pro pozitivní čtvercovou matici  $\mathbf{B}$  řádu  $n$  platí*

1.  $\rho(\mathbf{B}) \in \sigma(\mathbf{B})$ ,
2. *pokud je  $\mathbf{x} \neq \mathbf{0}$  vlastní vektor matice  $\mathbf{B}$  příslušný vlastnímu číslu  $\rho(\mathbf{B})$ , pak jsou buď všechny souřadnice vektoru  $\mathbf{x}$  kladné a nebo jsou všechny souřadnice vektoru  $\mathbf{x}$  záporné,*
3. *algebraická násobnost vlastního čísla  $\rho(\mathbf{B})$  je 1,*
4. *dimenze prostoru  $\mathcal{N}(\mathbf{B} - \rho(\mathbf{B})\mathbf{I}_n)$  se rovná 1, tj. každý nenulový vlastní vektor příslušný vlastnímu číslu  $\rho(\mathbf{B})$  je reálným násobkem libovolného jiného nenulového vlastního vektoru příslušného k témuž vlastnímu číslu.*

Z těchto vlastností vyplývá existence jednoznačně určeného *kladného* vlastního vektoru  $\mathbf{p} = (p_1, \dots, p_n)^T$  odpovídajícího spektrálnímu poloměru  $\rho(\mathbf{B})$ , pro který platí  $p_i > 0$  pro každé  $i = 1, \dots, n$  a dále

$$p_1 + p_2 + \dots + p_n = 1.$$

Tento jednoznačně určený kladný vlastní vektor nazýváme *Perronův vektor* kladné matice  $\mathbf{B}$  a vlastní číslo  $\rho(\mathbf{B})$  nazýváme *Perronův kořen* matice  $\mathbf{B}$ .

Perronův kořen a vektor nemůžeme přímo využít k výpočtu vektoru důležitosti webových stránek, protože matice incidence webu není kladná, je pouze nezáporná. Proto musíme použít následující zesílení předchozí věty, kterému se říká *Perronova-Frobeniova věta*.

**Věta 15.7** *Je-li  $\mathbf{B}$  nezáporná čtvercová matice řádu  $n$ , pro kterou platí, že  $\mathbf{B}^M$  je kladná matice pro dostatečně velký exponent  $M$ , pak platí*

1.  $\rho(\mathbf{B}) \in \sigma(\mathbf{B})$ ,
2. pokud je  $\mathbf{x} \neq \mathbf{0}$  vlastní vektor matice  $\mathbf{B}$  příslušný vlastnímu číslu  $\rho(\mathbf{B})$ , pak jsou buď všechny souřadnice vektoru  $\mathbf{x}$  kladné a nebo jsou všechny souřadnice vektoru  $\mathbf{x}$  záporné,
3. algebraická násobnost vlastního čísla  $\rho(\mathbf{B})$  je 1,
4. dimenze prostoru  $\mathcal{N}(\mathbf{B} - \rho(\mathbf{B})\mathbf{I}_n)$  se rovná 1, tj. každý nenulový vlastní vektor příslušný vlastnímu číslu  $\rho(\mathbf{B})$  je reálným násobkem libovolného jiného nenulového vlastního vektoru příslušného k témuž vlastnímu číslu.

Také v případě nezáporných matic také existuje jednoznačně určený vlastní (Perronův) vektor  $\mathbf{p} = (p_1, \dots, p_n)^T$  odpovídajícího spektrálnímu poloměru  $\rho(\mathbf{B})$ , pro který platí  $p_i > 0$  pro každé  $i = 1, \dots, n$  a dále

$$p_1 + p_2 + \dots + p_n = 1.$$

Matice incidence webu  $\mathbf{A}$  je sice nezáporná, s velkou pravděpodobností ale nespĺňuje druhý z předpokladů Perronovy-Frobeniovy věty, totiž že  $\mathbf{A}^M$  je kladná matice pro dostatečně velký exponent  $M$ . Toto už je ale drobný problém.

Tato idea porovnávání významu lidí nebo věcí na základě vlivu, který na sebe vzájemně mají, má mnoho nejrůznějších aplikací ve statistice i v dalších oborech. Anglicky se jí říká *Kendall-Wei theory of ranking* a pochází z padesátých let minulého století.

### Soustavy obyčejných lineárních diferenciálních rovnic

Soustavou obzčejných lineárních diferenciálních rovnic s konstatními koeficienty rozumíme soustavu rovnic následujícího tvaru.

$$u_1' = a_{11}u_1 + a_{12}u_2 + \dots + a_{1n}u_n,$$

$$\begin{aligned} u_2' &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n, \\ &\vdots \\ u_n' &= a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n, \end{aligned}$$

spolu s počátečními podmínkami

$$\begin{aligned} u_1(0) &= c_1, \\ u_2(0) &= c_2, \\ &\vdots \\ u_n(0) &= c_n. \end{aligned}$$

Předpokládáme, že matice  $\mathbf{A} = (a_{ij})$  je reálná matice, funkce  $u_i = u_i(t)$  jsou reálné funkce reálné proměnné a  $c_i$  jsou reálná čísla pro  $i = 1, \dots, n$ . Označíme-li  $\mathbf{c} = (c_1, \dots, c_n)^T$  a  $\mathbf{u} = (u_1(t), \dots, u_n(t))$ , můžeme soustavu zapsat v maticovém tvaru

$$\mathbf{u}' = \mathbf{A}\mathbf{u}, \quad \text{a} \quad \mathbf{u}(0) = \mathbf{c}.$$

V úvodu předchozí kapitoly jsme si řekli, že jedna obyčejná diferenciální rovnice  $u'(t) = \alpha u(t)$  s počáteční podmínkou  $u(0) = c$  má jednoznačně určené řešení  $u(t) = c \cdot e^{\alpha t}$ . Můžeme se proto pokusit najít analogický vzorec pro řešení soustavy o  $n$  neznámých funkcích.

Pokud je matice  $\mathbf{A}$  diagonalizovatelná a

$$\mathbf{A} = \lambda_1 \mathbf{E}_1 + \cdots + \lambda_k \mathbf{E}_k$$

je její spektrální rozklad podle Věty 10.15, můžeme se pokusit definovat matici funkcí reálné proměnné  $t$  jako

$$e^{\mathbf{A}t} = e^{\lambda_1 t} \mathbf{E}_1 + \cdots + e^{\lambda_k t} \mathbf{E}_k.$$

Matici funkcí reálné proměnné můžeme derivovat tak, že derivujeme každou funkci zvlášť. Takto definovaná matice reálných funkcí má následující vlastnosti.

**Tvrzení 15.8** *Je-li  $\mathbf{A}$  diagonalizovatelná matice řádu  $n$  a*

$$\mathbf{A} = \lambda_1 \mathbf{E}_1 + \cdots + \lambda_k \mathbf{E}_k$$

její spektrální rozklad, pak matice funkcí

$$e^{\mathbf{A}t} = e^{\lambda_1 t} \mathbf{E}_1 + \cdots + e^{\lambda_k t} \mathbf{E}_k.$$

má následující vlastnosti

$$\begin{aligned} \frac{d e^{\mathbf{A}t}}{dt} &= \mathbf{A} \cdot e^{\mathbf{A}t}, \\ \mathbf{A} \cdot e^{\mathbf{A}t} &= e^{\mathbf{A}t} \cdot \mathbf{A}, \\ e^{-\mathbf{A}t} e^{\mathbf{A}t} &= e^{\mathbf{A}t} e^{-\mathbf{A}t} = \mathbf{I}_n, \\ \mathbf{I}_n &= e^{\mathbf{0}}. \end{aligned}$$

Dále vektor funkcí  $\mathbf{u} = e^{\mathbf{A}t} \mathbf{c}$  je jediným řešením soustavy obyčejných lineárních rovnic  $\mathbf{u}' = \mathbf{A} \mathbf{u}$  s počátečními podmínkami  $u_i(0) = c_i$  pro  $i = 1, 2, \dots, n$ .

**Důkaz.** Abychom dokázali první rovnost, dosadíme za  $e^{\mathbf{A}t}$  podle definice. Dostaneme

$$\frac{d e^{\mathbf{A}t}}{dt} = \sum_{i=1}^k \lambda_i e^{\lambda_i t} \mathbf{E}_i = \left( \sum_{i=1}^k \lambda_i \mathbf{E}_i \right) \left( \sum_{i=1}^k e^{\lambda_i t} \mathbf{E}_i \right) = \mathbf{A} \cdot e^{\mathbf{A}t}.$$

V důkazu druhé rovnosti postupujeme podobně.

$$\begin{aligned} \mathbf{A} \cdot e^{\mathbf{A}t} &= \left( \sum_{i=1}^k \lambda_i \mathbf{E}_i \right) \left( \sum_{j=1}^k e^{\lambda_j t} \mathbf{E}_j \right) = \sum_{i=1}^k \lambda_i e^{\lambda_i t} \mathbf{E}_i \mathbf{E}_i = \\ &= \left( \sum_{j=1}^k e^{\lambda_j t} \mathbf{E}_j \right) \left( \sum_{i=1}^k \lambda_i \mathbf{E}_i \right) = e^{\mathbf{A}t} \cdot \mathbf{A}. \end{aligned}$$

Také třetí rovnost dokážeme analogicky.

$$\begin{aligned} e^{-\mathbf{A}t} e^{\mathbf{A}t} &= \left( \sum_{i=1}^k \lambda_i e^{-\lambda_i t} \mathbf{E}_i \right) \left( \sum_{j=1}^k \lambda_j e^{\lambda_j t} \mathbf{E}_j \right) = \sum_{i=1}^k e^{-\lambda_i t} e^{\lambda_i t} \mathbf{E}_i^2 = \\ &= \sum_{i=1}^k \mathbf{E}_i = \mathbf{I}_n = \left( \sum_{j=1}^k \lambda_j e^{\lambda_j t} \mathbf{E}_j \right) \left( \sum_{i=1}^k \lambda_i e^{-\lambda_i t} \mathbf{E}_i \right) = e^{\mathbf{A}t} e^{-\mathbf{A}t}. \end{aligned}$$

Poslední, čtvrtá rovnost vyplývá ihned přímo z definice  $e^{\mathbf{A}t}$  a vlastností spektrálních projektorů  $\mathbf{E}_i$ .

$$e^{\mathbf{0}} = \sum_{i=1}^k e^{\lambda_i \cdot 0} \mathbf{E}_i = \sum_{i=1}^k \mathbf{E}_i = \mathbf{I}_n.$$

Z první rovnosti ihned dostaneme, že vektor funkcí

$$\mathbf{u} = \mathbf{c} \cdot e^{\mathbf{A}t}$$

je řešením soustavy  $\mathbf{u}' = \mathbf{A}\mathbf{u}$ ,  $\mathbf{u}(0) = \mathbf{c}$ .

Je-li nějaký další vektor funkcí  $\mathbf{v}(t)$  řešením téže soustavy, tj. platí-li  $\mathbf{v}' = \mathbf{A}\mathbf{v}$  a  $\mathbf{v}(0) = \mathbf{c}$ , spočítáme napřed

$$e^{-\mathbf{A}\mathbf{v}} = \sum_{i=1}^k e^{-\lambda_i t} \mathbf{E}_i \mathbf{v}.$$

Odtud dostaneme

$$\begin{aligned} \frac{d(e^{-\mathbf{A}t}\mathbf{v})}{dt} &= -\sum_{i=1}^k \lambda_i e^{-\lambda_i t} \mathbf{E}_i \mathbf{v} + \sum_{i=1}^k e^{-\lambda_i t} \mathbf{E}_i \mathbf{v}' = -\mathbf{A} \cdot e^{-\mathbf{A}t}\mathbf{v} + e^{-\mathbf{A}t}\mathbf{v}' = \\ &= -e^{-\mathbf{A}t}\mathbf{A}\mathbf{v} + e^{-\mathbf{A}t}\mathbf{v}' = e^{-\mathbf{A}t}(-\mathbf{A}\mathbf{v} + \mathbf{v}') = \mathbf{0}. \end{aligned}$$

Odtud plyne, že všechny funkce  $v e^{-\mathbf{A}t}\mathbf{v}$  jsou konstantní. V bodě  $t = 0$  mají hodnoty

$$e^{\mathbf{0}}\mathbf{v}(0) = \mathbf{I}_n \mathbf{c} = \mathbf{c},$$

a tedy  $e^{-\mathbf{A}t}\mathbf{v} = \mathbf{c}$ . Vynásobíme poslední rovnost zleva maticí  $e^{\mathbf{A}t}$  a dostaneme

$$\mathbf{v} = e^{\mathbf{A}t} e^{-\mathbf{A}t}\mathbf{v} = e^{\mathbf{A}t}\mathbf{c} = \mathbf{u}.$$

Tím je dokázána jednoznačnost řešení  $\mathbf{u} = e^{\mathbf{A}t}\mathbf{c}$ .  $\square$

Mnoho fyzikálních nebo biologických procesů lze modelovat pomocí soustavy diferenciálních rovnic  $\mathbf{u}' = \mathbf{A}\mathbf{u}$ ,  $\mathbf{u}(0) = \mathbf{c}$ . Ukážeme si jeden příklad.

**Úloha 15.1** V čase  $t = 0$  vstříkneme do buňky číslo 1 jednotkové množství nějaké látky (léku). Látka se šíří z buňky číslo 1 do buňky číslo 2 buněčnou membránou. Předpokládáme, že rychlost šíření látky z buňky 1 do buňky 2 je přímo úměrná její koncentraci v buňce 1. Koeficient přímé úměrnosti označíme  $\alpha$ . Stejně tak předpokládáme, že rychlost šíření látky z buňky 2 do buňky 1 je přímo úměrná koncentraci této látky v buňce 2, koeficient přímé úměrnosti označíme  $\beta$ . Z fyzikálního významu koeficientů  $\alpha, \beta$  vyplývá předpoklad  $\alpha, \beta > 0$ . Určete koncentrace látky v obou buňkách v libovolném čase  $t > 0$ .

**Řešení.** Označíme koncentraci látky v první buňce v čase  $t$  jako  $u_1(t)$ . Podobně  $u_2(t)$  označuje koncentraci látky v druhé buňce. Z uvedených předpokladů vyplývají následující rovnice pro funkce  $u_1(t)$  a  $u_2(t)$

$$\begin{aligned} u_1'(t) &= \beta u_2(t) - \alpha u_1(t), \\ u_2'(t) &= \alpha u_1(t) - \beta u_2(t), \end{aligned}$$

a počáteční podmínky  $u_1(0) = 1$ ,  $u_2(0) = 0$ .

V maticovém tvaru soustavu zapíšeme jako  $\mathbf{u}' = \mathbf{A}\mathbf{u}$  a  $\mathbf{u}(0) = \mathbf{c}$ , kde

$$\mathbf{A} = \begin{pmatrix} -\alpha & \beta \\ \alpha & -\beta \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad \text{a} \quad \mathbf{c} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Podle předchozího Tvzení 15.8 můžeme zapsat řešení ve tvaru  $\mathbf{u} = e^{\mathbf{A}t}\mathbf{c}$ .

Matice  $\mathbf{A}$  má charakteristickou rovnici tvaru

$$(-\alpha - \lambda)(-\beta - \lambda) - \alpha\beta = \lambda^2 + (\alpha + \beta)\lambda = 0.$$

Vlastní hodnoty matice  $\mathbf{A}$  jsou proto  $\lambda_1 = 0$  a  $\lambda_2 = -(\alpha + \beta)$ . Protože jsou vlastní hodnoty různé, je matice  $\mathbf{A}$  diagonalizovatelná podle Tvzení 10.13.

Vlastní vektory příslušné vlastnímu číslu  $\lambda_1$  dostaneme jako řešení homogenní soustavy lineárních rovnic s maticí

$$\mathbf{A} - 0\mathbf{I}_2 = \mathbf{A} = \begin{pmatrix} -\alpha & \beta \\ \alpha & -\beta \end{pmatrix},$$

která má řešení ve tvaru násobků vektoru  $(\beta, \alpha)^T$ .

Podobně najdeme vlastní vektory příslušné vlastnímu číslu  $\lambda_2 = -(\alpha + \beta)$  jako řešení homogenní soustavy lineárních rovnic s maticí

$$\mathbf{A} + (\alpha + \beta)\mathbf{I}_2 = \begin{pmatrix} \beta & \beta \\ \alpha & \alpha \end{pmatrix},$$

která má řešení ve tvaru násobků vektoru  $(1, -1)^T$ .

Podle důkazu Spektrální věty pro diagonalizovatelné matice 10.15 můžeme k výpočtu spektrálních projektorů  $\mathbf{E}_1$  a  $\mathbf{E}_2$  použít regulární matici

$$\mathbf{P} = \begin{pmatrix} \beta & 1 \\ \alpha & -1 \end{pmatrix}.$$

Spočítáme inverzní matici

$$\mathbf{P}^{-1} = \begin{pmatrix} \frac{1}{\alpha + \beta} & \frac{1}{\alpha + \beta} \\ \frac{\alpha}{\alpha + \beta} & -\frac{\beta}{\alpha + \beta} \end{pmatrix}$$

a matice

$$\mathbf{E}_1 = \mathbf{P} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{P}^{-1} = \begin{pmatrix} \frac{\beta}{\alpha + \beta} & \frac{\beta}{\alpha + \beta} \\ \frac{\alpha}{\alpha + \beta} & \frac{\alpha}{\alpha + \beta} \end{pmatrix} = \frac{1}{\alpha + \beta} \begin{pmatrix} \beta & \beta \\ \alpha & \alpha \end{pmatrix}.$$

Podobně

$$\mathbf{E}_2 = \mathbf{P} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{P}^{-1} = \frac{1}{\alpha + \beta} \begin{pmatrix} \alpha & -\beta \\ -\alpha & \beta \end{pmatrix}.$$

Proto

$$e^{\mathbf{A}t} = \frac{1}{\alpha + \beta} \left[ e^{0t} \begin{pmatrix} \beta & \beta \\ \alpha & \alpha \end{pmatrix} + e^{-(\alpha+\beta)t} \begin{pmatrix} \alpha & -\beta \\ -\alpha & \beta \end{pmatrix} \right]$$

a tedy

$$\begin{aligned} \mathbf{u} = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} &= \frac{1}{\alpha + \beta} \left[ \begin{pmatrix} \beta & \beta \\ \alpha & \alpha \end{pmatrix} + e^{-(\alpha+\beta)t} \begin{pmatrix} \alpha & -\beta \\ -\alpha & \beta \end{pmatrix} \right] \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{\beta}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta} e^{-(\alpha+\beta)t} \\ \frac{\alpha}{\alpha + \beta} - \frac{\alpha}{\alpha + \beta} e^{-(\alpha+\beta)t} \end{pmatrix}. \end{aligned}$$

Limita pro  $t \rightarrow \infty$  se pak rovná

$$\lim_{t \rightarrow \infty} u_1(t) = \frac{\beta}{\alpha + \beta}, \quad \lim_{t \rightarrow \infty} u_2(t) = \frac{\alpha}{\alpha + \beta}.$$

□

### Filtry odstraňující šum

V závěru předchozí kapitoly jsme si ukázali rozklad matice  $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ , kde  $\mathbf{U}$  a  $\mathbf{V}$  jsou ortogonální matice, a

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_{r \times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

pro regulární matici  $\mathbf{D}_{r \times r}$  řádu  $r$ . Tento rozklad je užitečný při odstraňování šumu v nějakých datech. Předpokládáme, že naše data jsou v podobě matice  $\mathbf{A}$  tvaru  $m \times n$ . Napišeme si její rozklad

$$\mathbf{A} = \mathbf{U} \begin{pmatrix} \mathbf{D}_{r \times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}^T = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T,$$

kde  $\mathbf{u}_i$  je  $i$ -tý sloupec ortogonální matice  $\mathbf{U}$  řádu  $m$  a  $\mathbf{v}_j$  je  $j$ -tý sloupec ortogonální matice  $\mathbf{V}$  řádu  $n$  a  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$  jsou singulární hodnoty matice  $\mathbf{A}$ .

Označme si  $\mathbf{u}_i \mathbf{v}_i^T = \mathbf{Z}_i$  pro  $i = 1, \dots, r$ . Podle Příkladu 8.8 předpis

$$\langle \mathbf{B} | \mathbf{C} \rangle = \text{tr}(\mathbf{B}^T \mathbf{C})$$

definuje skalární součin na prostoru  $\mathcal{R}^{m \times n}$  reálných matic tvaru  $m \times n$ .

Spočítáme, že

$$\langle \mathbf{Z}_i | \mathbf{Z}_i \rangle = 1$$

pro  $i = 1, 2, \dots, r$  a

$$\langle \mathbf{Z}_i | \mathbf{Z}_j \rangle = 0,$$

pokud  $i \neq j$ . To znamená, že posloupnost matic  $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_r$  je ortonormální a vyjádření

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \mathbf{Z}_i$$

je Fourierův rozklad matice  $\mathbf{A}$  podle Definice 8.15.

Čím větší je singulární hodnota  $\sigma_i$ , tím větší část našich dat je ve “směru” matice  $\mathbf{Z}_i$ . Pokud předpokládáme, že šum je v datech obsažen “náhodně”, nezávisle na směru, můžeme předpokládat, že v každém směru je přibližně stejné množství šumu. Zvolíme nějaké  $k$ . Potom ve výrazu

$$\sigma_{k+1} \mathbf{Z}_{k+1} + \dots + \sigma_r \mathbf{Z}_r$$

je obsažena poměrně malá část dat a větší část šumu, zatímco ve výrazu

$$\sigma_1 \mathbf{Z}_1 + \dots + \sigma_k \mathbf{Z}_k$$

je obsažena mnohem větší část dat než šumu. To vysvětluje, že pokud místo matice

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \mathbf{Z}_i$$

pracujeme dále s maticí

$$\mathbf{A}' = \sum_{i=1}^k \sigma_i \mathbf{Z}_i$$

bude mít nová matice  $\mathbf{A}'$  větší poměr data/šum než původní matice  $\mathbf{A}$ .

Volba vhodného  $k$  závisí na konkrétní aplikaci. Dobrým počátečním krokem je zvolit takové  $k$ , aby rozdíl  $\sigma_k - \sigma_{k+1}$  byl co největší.

Tato myšlenka odfiltrování šumu z dat se používá například v některých vyhledávacích, ve filtrech odstraňujících spamy, atd.

### Vyhledávače

Představme si, že máme posloupnost slov (termínů)  $T_1, \dots, T_m$ , jejichž výskyt v dokumentech  $D_1, \dots, D_n$  nás zajímá. Každý dokument  $D_j$  prohlédneme a najdeme frekvenci (počet výskytů)  $f_{ij}$  každého termínu  $T_i$  v

dokumentu  $D_j$ . Získáme tak *frekvenční matici*  $\mathbf{F} = (f_{ij})$  tvaru  $m \times n$ . Každému dokumentu  $D_j$  tak rovněž přiřadíme vektor  $\mathbf{f}_j = \mathbf{F}_{*j}$ , sloupcový vektor matice  $\mathbf{F}$ . Většina prvků matice  $\mathbf{F}$  se bude rovnat 0.

Nzní položíme dotaz  $\mathbf{q} = (q_1, \dots, q_n)^T$ , kde  $q_i = 1$ , pokud nás zajímá termín  $T_i$  a  $q_i = 0$ , pokud nás termín  $T_i$  nezajímá. Jako míru toho, jak se vektor (dotaz)  $\mathbf{q}$  liší od frekvenčního vektoru  $\mathbf{f}_j$  dokumentu  $T_j$  můžeme zvolit standardní skalární součin

$$\cos \alpha_j = \frac{\mathbf{q}^T \mathbf{f}_j}{\|\mathbf{q}\| \cdot \|\mathbf{d}_j\|} = \frac{\mathbf{q}^T \mathbf{F} \mathbf{e}_j}{\|\mathbf{q}\| \cdot \|\mathbf{F} \mathbf{e}_j\|},$$

kde  $\mathbf{e}_j$  je  $j$ -tý vektor standardní báze prostoru  $\mathcal{R}^{n \times 1}$ . Pokud je  $|\cos \alpha_j| \geq \tau$  pro nějakou zvolenou hranici  $\tau$ , tak dokument  $D_j$  považujeme za relevantní a zařadíme jej do odpovědi na dotaz  $\mathbf{q}$ . Volba vhodného  $\tau$  je záležitostí zkušenosti a umění.

Pokud jsou sloupce matice  $\mathbf{F}$  předem normalizovány, tj. mají délku 1, pak vektor

$$\mathbf{q}^T \mathbf{F} = (\cos \alpha_1, \cos \alpha_2, \dots, \cos \alpha_n)$$

umožňuje posoudit relevanci každého dokumentu  $D_j$ .

Musíme ale předpokládat, že frekvenční matice  $\mathbf{F}$  je zatížena šumem, který pochází z překlepů v dokumentech, nejednoznačnosti v používání jazyka, atd. Souřadnice vektoru  $\mathbf{q}^T \mathbf{F}$  proto nejsou příliš spolehlivou mírou relevance dokumentů pro dotaz  $\mathbf{q}$ . Použijeme proto předchozí metodu odstraňování šuku pomocí SDV-rozkladu. Napíšeme si matici  $\mathbf{F} = \mathbf{U} \mathbf{D} \mathbf{V}^T$  ve tvaru

$$\mathbf{F} = \mathbf{U} \mathbf{D} \mathbf{V}^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T,$$

a místo matice  $\mathbf{F}$  budeme dále pracovat pouze s maticí

$$\mathbf{F}_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T.$$

Místo čísel  $\cos \alpha_j$  tak dostaneme čísla

$$\cos \beta_j = \frac{\mathbf{q}^T \mathbf{F}_k \mathbf{e}_j}{\|\mathbf{q}\| \cdot \|\mathbf{F}_k \mathbf{e}_j\|}.$$