Cvičení z optimalizace:
Rozklad matice podle singulárních čísel (SVD)

Vojtěch Franc, Tomáš Werner

2011

Úloha 1: Aproximace bodů přímkou

Je dáno m bodů v rovině $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m \in \mathbb{R}^2$. Najděte přímku v rovině (tedy afinní podprostor dimenze 1 prostoru $\mathbb{R}^2$) takovou, aby součet čtverců *kolmých* vzdáleností bodů k této přímce byl minimální. Představte si např., že někdo body naklikal myší v grafickém rozhraní a vaším úkolem je proložit jimi nejlepší přímku.

Nápověda: Optimální přímka prochází těžištěm bodů.

Úkoly k vypracování

1. Stáhněte si soubor

   `http://cmp.felk.cvut.cz/cmp/courses/OPT/cviceni/mocap/data_A.mat`.

   Ten obsahuje matici $\mathbf{A} \in \mathbb{R}^{m \times 2}$ s řádky $\boldsymbol{a}_1^T, \dots, \boldsymbol{a}_m^T$.

2. Zobrazte do jednoho obrázku zadané body $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ (modře) a jejich kolmé projekce $\tilde{\boldsymbol{a}}_1, \dots, \tilde{\boldsymbol{a}}_m$ na nalezenou přímku (červeně). Popište stručně, jak jste výsledné body nalezli.

3. Jaký je součet čtverců kolmých vzdáleností bodů k nalezené přímce?

4. Nalezněte požadovanou přímku ve dvou různých reprezentacích:

   $$\{\, \boldsymbol{y} \in \mathbb{R}^2 \mid \boldsymbol{y}^T \boldsymbol{x} = \alpha \,\} = \{\, \boldsymbol{y}_0 + t\boldsymbol{s} \mid t \in \mathbb{R} \,\}.$$

   Výstupem úkolu budou $\boldsymbol{x} \in \mathbb{R}^2$, $\alpha \in \mathbb{R}$, $\boldsymbol{y}_0 \in \mathbb{R}^2$, $\boldsymbol{s} \in \mathbb{R}^2$. Vektory zvolte tak, aby $\|\boldsymbol{x}\| = 1$, $\|\boldsymbol{s}\| = 1$, a aby norma $\|\boldsymbol{y}_0\|$ byla minimální (tj. bod $\boldsymbol{y}_0$ byl kolmý průmět počátku $\mathbf{0}$ na přímku).

Úloha 2: Komprese sekvence z motion capture

Při tvorbě počítačových her nebo animovaných filmů se používá technologie *motion capture* (pokud víme, termín nemá zavedený český ekvivalent). Na živého herce se připevní terčíky odrážející infračervené světlo. Terčíky se připevňují na významné body na těle, jako klouby apod. Množství speciálních kamer pak snímá polohy terčíků a z těch se počítá poloha každého terčíku v třírozměrném prostoru pro každý snímek. Polohy terčíků v prostoru se pak použijí např. pro animaci postav syntetizovaných počítačovou grafikou. Viz např. Wikipedie.

Pro získání plynulého pohybu je třeba snímat s vysokou frekvencí. Například data použitá v naší úloze byla snímána s vzorkovací frekvencí 120 Hz. Ve výsledku je pak třeba pracovat s velkými objemy dat, což je nevýhoda. Naším úkolem bude snížit objem dat tak, abychom ztratili co nejméně užitečné informace.

Prostorová poloha jednoho terčíku v jednom snímku je dána trojicí souřadnic. V našem případě máme $\ell = 41$ terčíků. Poloha všech terčíků v jednom snímku je dána vektorem $\boldsymbol{a} \in \mathbb{R}^n$ délky $n = 3\ell$. Ten si lze představit jako bod v n-rozměrném prostoru. Celkově máme m snímků, tedy vektory $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in \mathbb{R}^n$.

Hledáme body, které co nejlépe aproximují původní body a zároveň se dají reprezentovat menším objemem dat. Přesněji, hledáme body $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_m$, které leží v afinním podprostoru dimenze $r < n$ lineárního prostoru $\mathbb{R}^n$ a zároveň minimalizují součet čtverců vzdáleností od původních bodů, tj. číslo

$$\sum_{i=1}^{m} \|\tilde{\boldsymbol{a}}_i - \boldsymbol{a}_i\|^2. \tag{1}$$

Z přednášek víte, že optimální afinní podprostor prochází těžištěm bodů

$$\boldsymbol{c} = \frac{1}{m} \sum_{i=1}^{m} \boldsymbol{a}_i.$$

Když tedy body posunete tak, aby jejich těžiště leželo v počátku, tak místo afinního podprostoru můžete hledat podprostor lineární.

Jak dochází ke kompresi dat? Na první pohled se zdá, že původní i aproximované body zabírají stejné množství paměti ($m \times n$ čísel), takže jsme nic neušetřili. Ovšem nové body leží v podprostoru nižší dimenze r, který lze reprezentovat bází. Nechť tato báze je ortonormální a tvoří sloupce matice $\tilde{\mathbf{V}} \in \mathbb{R}^{n \times r}$. Pak každý bod v podprostoru je lineární kombinace bázových vektorů,

$$\tilde{\boldsymbol{a}} = \tilde{\mathbf{V}} \boldsymbol{y}, \tag{2}$$

kde složky vektoru $\boldsymbol{y} \in \mathbb{R}^r$ jsou koeficienty lineární kombinace. Pro uložení všech bodů $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_m$ tedy musíme uložit bázové vektory ($r \times n$ čísel, což je zanedbatelné) a pak vektory $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ ($r \times m$ čísel).

Úkoly k vypracování

1. Stáhněte si soubory `makarena1.txt` a `walk1.txt` z adresáře[1]

   `http://cmp.felk.cvut.cz/cmp/courses/OPT/cviceni/mocap/`.

   Tyto soubory popisují sekvence 'Tanec Makarena' a 'Chůze'. Každý soubor obsahuje matici rozměru $m \times n$ (označte ji $\mathbf{A}$) s řádky $\boldsymbol{a}_1^T, \ldots, \boldsymbol{a}_m^T$. Pro vizualizaci sekvencí použijte skript `show_motion.m`.

2. Minimalizujte kritérium (1) za podmínky, že body $\tilde{\boldsymbol{a}}_1, \ldots, \tilde{\boldsymbol{a}}_m$ leží v afinním podprostoru dimenze r. Výsledkem bude matice $\tilde{\mathbf{A}}$ s řádky $\tilde{\boldsymbol{a}}_1^T, \ldots, \tilde{\boldsymbol{a}}_m^T$.

[1] Sekvence byly staženy ze stránek `http://movlab.ulusofona.pt`.

Proveďte pro sekvenci 'Chůze' a pro pět různých hodností $r \in \{1, 2, 5, 10, 15\}$. Výsledkem je tabulky s optimálními hodnotami kritéria (1) pro všechny tyto hodnosti.

Pro kontrolu si jednotlivá řešení přehrajte skriptem `show_motion2.m`, který zobrazí sekvenci v matici $\mathbf{A}$ spolu se sekvencí v matici $\tilde{\mathbf{A}}$. Pokud je vaše řešení správné, budou obě sekvence podobné.

3. Výsledné body vyjádřete jako lineární kombinaci (2) bázových vektorů. Pro $r = 2$ nakreslete sekvenci vektorů $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ jako trajektorii v rovině (po sobě jdoucí body spojte úsečkou). Proveďte pro sekvence 'Tanec Makarena' a 'Chůze'. Výsledkem budou tedy dva obrázky.

   Zkuste si nakreslit tuto trajektorii i ve třírozměrném prostoru, tedy pro $r = 3$. Pro 'Chůzi' je aproximace dosti dobrá, a přitom každý snímek sekvence je reprezentován pouhými třemi (místo $3\ell = 123$) čísly!

4. Uvažujte že postava dělá čistý translační pohyb, tj. konfigurace bodů se nemění a jejich souřadnice se pohybují po přímce. Jaká je minimální dimenze podprostoru, aby aproximační chyba byla nulová? Výsledkem je teoretické odůvodnění.

5. Dejme tomu, že bychom chtěli spočítat optimální chyba aproximace (1) pro různé hodnosti $r \leq n$. Dostali bychom tedy n čísel, z nichž poslední by bylo nulové. Jaký vztah mají tato čísla k singulárním číslům? Výsledkem je vzorec. Jako inspiraci si vykreslete singulární čísla seřazená do grafu.