Algorithms for Networked Information
DM2

November 20, 2015

Exercise 1 Singular values and matrix completion

We recall the matrix norms $\|A\|_2 = \sup_{x\neq 0}\|Ax\|_2/\|x\|_2$ and $\|A\|_F = \left(\sum_{i,j}|A_{i,j}|^2\right)^{1/2}$.

1. Prove that $\|A\|_2 \le \|A\|_F \le \sqrt{r}\|A\|_2$ where r is the rank of A.

We denote the SVD of A by $A = \sum_{i=1}^{\min(n,m)} \sigma_i(A)\mathbf{u}_i\mathbf{v}_i^T$. For a matrix A, we define A_k a *rang k approximation* of A by $A_k = \sum_{i=1}^{k}\sigma_i(A)\mathbf{u}_i\mathbf{v}_i^T$.

2. For any $i \le \min(n,m)$ and vectors $\mathbf{w_1},\dots,\mathbf{w_{i-1}}$, show that

$$\sigma_i(A) \le \max_{\mathbf{v}\perp\mathbf{w_1},\dots,\mathbf{w_{i-1}}} \frac{\|Av\|_2}{\|v\|_2}.$$

3. Let A,B be real matrices of dimension $m\times n$. Show that if $i+j \le 1+\min\{m,n\}$, then

$$\sigma_{i+j-1}(A) \le \sigma_i(B) + \sigma_j(A-B).$$

   Deduce that

$$\max_{1\le k\le m\wedge n} |\sigma_k(A)-\sigma_k(B)| \le \|A-B\|_2.$$

4. Show that

$$\|B-B_k\|_2 \le \|A-A_k\|_2 + \|A-B\|_2.$$

   Deduce that $\|A-B_k\|_2 \le \|A-A_k\|_2 + 2\|A-B\|_2$.

We consider the following scenario: Let M be a $m\times n$ matrix with $m\le n$ of rank r (modeling user rankings). Let $p\in[0,1]$. We assume that each entry of M is observed with probability p and not observed with probability $1-p$, independently of the other entries. For simplicity, we assume $M_{ij}\in[0,1]$ and that p and r are known. We construct an estimate $\hat{M}$ of M starting from the observed entries as follows:

a- Let X be the matrix with $x_{ij} = m_{ij}$ if the entry is observed and $x_{ij} = 0$ else. Let $X = \sum_{i=1}^{m} s_i\mathbf{u}_i{\mathbf{v}_i}^T$ be its singular value decomposition.

b- We define

$$W = \frac{1}{p}\sum_{i\le r} s_i\mathbf{u}_i{\mathbf{v}_i}^T.$$

and matrix $\hat{M}$ by $\hat{m}_{ij} = m_{ij}$ if the entry is observed and else by

$$\hat{m}_{ij} = \begin{cases} w_{ij} & \text{if } 0\le w_{ij}\le 1,\\ 1 & \text{if } w_{ij} > 1,\\ 0 & \text{if } w_{ij} < 0.\end{cases}$$

We define the mean squared error

$$\mathrm{MSE}(\hat{M}) = \mathbb{E}\left[\frac{1}{mn}\|M-\hat{M}\|_F^2\right].$$

5. Show that $\|M-W\|_F^2 \leq 8r\|M-\frac{1}{p}X\|_2^2$. Deduce that

$$\mathrm{MSE}(\hat{M}) \leq \frac{8r}{mnp^2}\mathbb{E}\left[\|pM-X\|_2^2\right]$$

To bound the last term, we use the matrix version of the Bernstein inequality: Let $Z_1,\dots,Z_k \in \mathbb{R}^{m\times n}$ be random symmetric matrices such that $\mathbb{E}[Z_t]=0$, $\|Z_t\|_2 \leq 1$ and $\max\left\{\|\sum_{t=1}^k \mathbb{E}[Z_t Z_t^T]\|_2; \|\sum_{t=1}^k \mathbb{E}[Z_t^T Z_t]\|_2\right\} \leq \sigma^2$. Then

$$\mathbb{P}\left(\|\sum_{t=1}^k Z_t\|_2 \geq s\right) \leq (m+n)\exp\left(-\frac{s^2}{2(\sigma^2+s/3)}\right).$$

6. We define $Y = pM - X$. Find $\mathbb{E}[Y]$. Show that $\|Y\|_2^2 \leq nm$.
7. Show that
$$\mathbb{P}(\|Y\| \geq s) \leq (m+n)\exp\left(-\frac{s^2}{2(np+s/3)}\right)$$
8. Show that $\mathbb{E}\left[\|Y\|_2^2\right] \leq nm\mathbb{P}(\|Y\|_2 \geq s) + s^2$. Deduce that for $np > \ln n$, there is a constant C such that:
$$\mathrm{MSE}(\hat{M}) \leq \frac{Cr\log n}{mp}.$$
9. For $m = 100$, $n = 1000$ and $r = 10$, generate a matrix M (you can instead use the one available on my website). For various values of $p \in [0,1]$, compute an empiric esitmate of $MSE(\hat{M})$ and plot the corresponding picture.