Application: Matrix Completion

Note Title 2015-11-20

M: Fixed $n \times n$ matrix w/ $\text{rank}(M) = r \ll n$.

Problem | Given a random subset of entries, how to estimate missing entries?

Applications: Recommender systems, data imputation, sensor triangulation...

Assume:

① Each entry is viewed independently w/ prob. p.

② M is spread i.e. $\|M\|_\infty := \max_{i,j} |M_{ij}| \leq 1$

for simplicity

Let $\cdot\ \Omega = \{\text{viewed entries}\} \subseteq \{1,\dots n\} \times \{1,\dots n\}$

$\cdot\ M_\Omega := \begin{cases} M_{ij} & (i,j) \in \Omega \\ 0 & (i,j) \notin \Omega \end{cases}$

Observe: $\mathbb{E} \underbrace{\frac{1}{p} M_\Omega}_{\text{unbiased estimator}} = M$

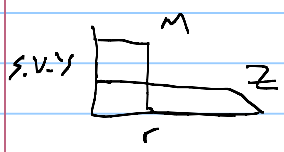
but $\frac{1}{p} M_\Omega$ is a terrible estimate of M!

Write $\frac{1}{p} M_\Omega = M + Z$ $\quad Z := \frac{1}{p} M_\Omega - M$

rank-r matrix $\quad$ random matrix

Estimate M: $\hat{M} = P(M+Z) = \underset{\text{rank}(M') \leq r}{\text{argmin}} \|M' - (M+Z)\|_F$

projection onto set of rank-r matrices



How large is error $\|\hat{M}-M\|_F$?

Observe:

① $\text{rank}(\hat{M}-M) \le 2r$

② $\|\hat{M}-M\| \le \|\hat{M}-\frac{1}{p}M_\Omega\| + \|M-\frac{1}{p}M_\Omega\|$

$\le 2\|M-\frac{1}{p}M_\Omega\|$ (Check this. Takes a bit of work.)

$= 2\|Z\|$

① + ② $\Rightarrow$

$$\|\hat{M}-M\|_F^2 = \sum_{i=1}^{2r} \sigma_i(\hat{M}-M)^2 \le \underbrace{\|\hat{M}-M\|}_{\text{max singular value}}{}^2 \cdot 2r = 8r\|Z\|^2 \quad (*)$$

$\Rightarrow$ To bound error, we only need to bound $\|Z\|$.

Case 1: $p=\frac{1}{2} \Rightarrow Z_{ij} = \begin{cases} M_{ij} & \text{w.p. } \frac{1}{2} \\ -M_{ij} & \text{w.p. } \frac{1}{2} \end{cases}$.

$|M_{ij}| \le 1$ by assumption.

$\Rightarrow \|Z_{ij}\|_{\psi_2} \le 1$

$\Rightarrow \|Z\|^2 \le cn$ w.h.p.

(By many of the methods from class.)

Thus, using $(*)$,

avg error per entry $\left\{ \dfrac{\|\hat{M}-M\|_F^2}{n^2} \le c\,\dfrac{r}{n} \right.$ w.h.p.

Case 2. $p \ll \frac{1}{2}$. Use matrix-Bernstein ≠.

Thm (General matrix Bernstein ≠)

Let $X_1, \dots X_m$ be ind, mean-zero, $n_1 \times n_2$, random matrices. Suppose

$$\|X_i\| \le k \text{ a.s.}, \quad \max\{\|\sum_i \mathbb{E} X_i X_i^*\|, \|\sum_i \mathbb{E} X_i^* X_i\|\} \le \sigma^2.$$

Then,

$$P(\|\sum_{i=1}^m X_i\| \ge t) \le (n_1+n_2) \exp\left(\frac{-t^2/2}{\sigma^2 + \frac{kt}{3}}\right)$$

Control k, σ^2.

k: range; σ^2: variance

k

$$Z = \sum_{i,j} z_{ij} e_i e_j = \sum_{i,j} X_{ij}$$

e_i: ith basis vector

$$z_{ij} = \begin{cases} (\frac{1}{p}-1) M_{ij} & \text{w.p. } p \\ -M_{ij} & \text{w.p. } 1-p \end{cases}$$

$$\Rightarrow \|X_{ij}\| \le \frac{1}{p}\|M\|_\infty = \frac{1}{p} =: k$$

σ^2

$$\mathbb{E} X_{ij} X_{ij}^* = \mathbb{E}(z_{ij})^2 \cdot e_i e_i^T = \frac{p(1-p)|M_{ij}|^2}{p^2} e_i e_i^T$$
$$\preceq \frac{1}{p} e_i e_i^T$$

$$\Rightarrow \sum_{i,j} \mathbb{E} X_{ij} X_{ij}^* \preceq \sum_j \sum_i p e_i e_i^T = \sum_j pI = \frac{n}{p} I$$

Similarly for $\sum_{i,j} \mathbb{E} X_{ij}^* X_{ij}$

$$\Rightarrow \sigma^2 := \frac{n}{p}$$

Matrix Bernstein I

$$\Rightarrow \quad P(\|Z\| > t) \le 2n \exp\left(\frac{-t^2/2}{\frac{n}{p} + \frac{t}{3 \cdot p}} \right)$$

$$\Rightarrow \|Z\| \le C\sqrt{\frac{n}{p} \log n}$$

(*) gives, w.h.p,

$$\frac{\|\hat{M} - M\|_F^2}{n^2} \le C \frac{r}{pn} \log(n) = C \frac{rn}{m} \log(n)$$

where $m = pn^2 = \mathbb{E}[\#$ of viewed entries$]$