MATRIX COMPLETION

Setting:

$M \in \mathbb{R}^{D \times N}$, $\text{rank}(M) = d << D$.

Ω = observed indexes

$= \{(i,j) : M_{i,j} \text{ is } \underline{\underline{\text{known}}}\}$

$P_\Omega : \mathbb{R}^{D \times N} \to \mathbb{R}^{D \times N}$ = "projection" on Ω

$$P_\Omega(X) = \begin{cases} X_{ij} & (i,j) \in \Omega \\ 0 & (i,j) \notin \Omega. \end{cases}$$

OPTIMISATION PROBLEM #1:

$$\min_X \text{rank}(X) \quad \text{s.t.} \quad P_\Omega(X) = P_\Omega(M)$$

↳ Find X such that:

(1) X agrees with M in Ω.

(2) X has lowest possible rank

Properties:

(1) M - non-sparse, Ω = "randomly distributed"

→ unique solution to #1.

(2) Solution is NP-Hard. (very difficult to solve)

OPTIMISATION PROBLEM #2:

$$\min_X \|X\|_* \quad \text{s.t.} \quad P_\Omega(X) = P_\Omega(M)$$

nuclear norm $= \sum_{i=1}^{r} \sigma_i$

Properties:

ild assumptions $\Rightarrow$ #1 = #2.

he is an algorithm, BUT generally slow.

OPTIMISATION PROBLEM #3:

$$\min_X \tau\|X\|_* + \frac{1}{2}\|X\|_F^2 \quad \text{s.t.} \quad P_\Omega(X) = P_\Omega(M)$$

Properties:

(1) Simple algorithm

(2) $\tau \to \infty \Rightarrow$ #3 converges to #2.

Largragian:

want to be 0.

$$\mathcal{L}(X, \Lambda) = \tau\|X\|_* + \frac{1}{2}\|X\|_F^2 + \langle \Lambda, P_\Omega(M) - P_\Omega(X) \rangle$$

Iterations:

(1) $X^{(k+1)} = \operatorname{argmin}_X \mathcal{L}(X, \Lambda^{(k)})$

(2) $\Lambda^{(k+1)} = \Lambda^{(k)} + \beta\left(P_\Omega(M) - P_\Omega(X^{(k+1)})\right)$

Step (2):

$$\underset{X}{\operatorname{argmin}}\ \tau\|X\|_* + \frac{1}{2}\|X\|_F^2 - \langle \Lambda, \underbrace{P_\Omega(X)}_{\langle P_\Omega(\Lambda),\ X\rangle} \rangle$$

exercise.

$$= \underset{X}{\operatorname{argmin}}\ \tau\|X\|_* + \frac{1}{2}\|X - P_\Omega(\Lambda)\|_F^2$$

Solution: $X^* = D_\tau\left(P_\Omega(\Lambda)\right)$ $\boxed{U\cdot\Sigma'\cdot V^T}$

$$= \sum_{i=1}^{r} S_\tau(\sigma_i)\, u_i\, v_i^T = U\cdot S_\tau(\Sigma')\cdot V^T$$

ALGORITHM

Set: $\Lambda^{(0)} = 0$

Iterate: $X^{(k+1)} = D_\tau\left(P_\Omega(\Lambda^{(k)})\right)$

$$\Lambda^{(k+1)} = \Lambda^{(k)} + \beta P_\Omega\left(M - X^{(k+1)}\right)$$

Return $X^{(k)}$

Properties:

(1) Simple.

(2) Converges to #3.

(3) τ - "large enough" $\to$ converges to #2.

(4) Output has low rank.

(5) $\Lambda^{(k)}$ is always zero outside Ω.

(6) Sufficient to find SVD only for $\sigma_i > \tau$. (u_i, v_i)