Robust PCA

Assumption: $X = L_0 + E_0$

L_0: low rank

E_0: sparse (outliers)

OPTIMISATION PROBLEM:

$$\circledast \quad \min_{L,E} \left(\operatorname{rank}(L) + \lambda \|E\|_0 \right) \quad \text{s.t.} \quad X = L + E$$

$\|E\|_0$: # of non-zero values

Some issues:

(1) Solution non-unique:

ex:

$$X = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & 0 & \\ 0 & & & \end{pmatrix}$$

low rank

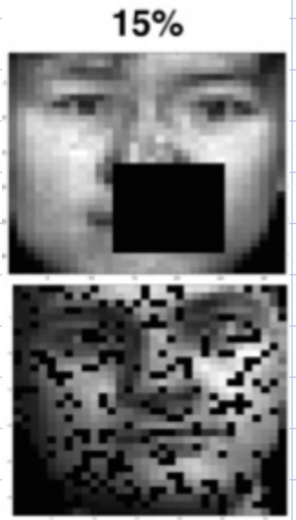
sparse

We can take: $L = X, \ E = 0$

$E = X, \ L = 0$

$L = \frac{1}{2}X, \ E = \frac{1}{2}X.$

(2) Outliers with "patterns":



impossible

possible to recover

ASSUMPTIONS:

(1) L_0 - is not sparse

(2) Outliers are uniformly distributed within the matrix.

⊛ non convex, non-differentiable

$\boxed{\text{HARD}}$

RELAXED OPTIMISATION PROBLEM

Nuclear norm: M - a matrix.

$\sigma_1, \_\_\_, \sigma_R$ - singular values.

$$\|M\|_* = \sum_{i=1}^{R} \sigma_i$$

$$\min_{L,E} \|L\|_* + \lambda \|E\|_1 \quad \text{s.t.} \quad X = L + E.$$

NAME: PRINCIPAL COMPONENT PURSUIT (PCP)

⊛ CANDÉS et. al. 2011 - conditions for exact solution.

ALM = AUGMENTED LAGRANGE MULTIPLIER

$$\min_{L,E} \|L\|_* + \alpha\|E\|_1 \quad \text{s.t.} \quad X = L + E$$

$X - L - E = 0$

$$\mathcal{L}(L, E, \Lambda) = \|L\|_* + \alpha\|E\|_1$$

$$+ \underbrace{\langle \Lambda, X - L - E \rangle}_{\sum_{i,j} \Lambda_{ij} (X-L-E)_{ij}} + \frac{\beta}{2}\|X-L-E\|_F^2$$

(1) EXACT ALM METHOD:

Iterations:

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

(2) $\Lambda^{(k+1)} = \Lambda^{(k)} + \beta(X - L^{(k+1)} - E^{(k+1)})$

(2) ALTERNATING DIRECTION METHOD of MULTIPLIERS (ADMM)

Break (1) into 2 separate steps:

(i) $L^{(k+1)} = \operatorname{argmin}_L \mathcal{L}(L, E^{(k)}, \Lambda^{(k)})$

(ii) $E^{(k+1)} = \operatorname{argmin}_E \mathcal{L}(L^{(k+1)}, E, \Lambda^{(k)})$

(iii) $\Lambda^{(k+1)} = \Lambda^{(k)} + \beta(X - L^{(k+1)} - E^{(k+1)})$

DETOUR - OPTIMISATION

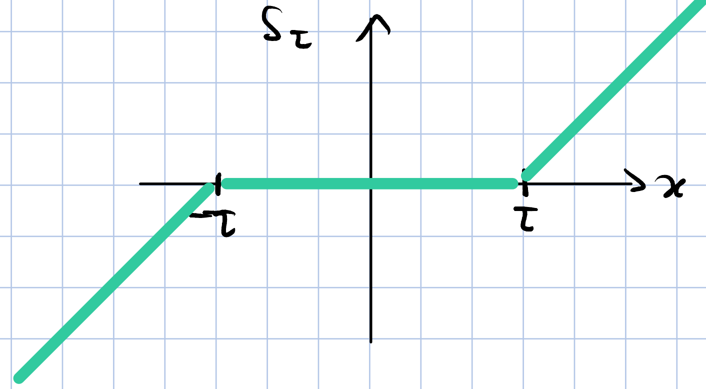
Suppose $\underline{x} \in \mathbb{R}^n$.

(#) $\underline{a}^* = \underset{\underline{a}}{\operatorname{argmin}} \; \frac{1}{2}\|\underline{x}-\underline{a}\|_2^2 + \tau\|\underline{a}\|_1$

looking for sparse approx. for $\underline{x}$

$\underline{a}^* = \left(S_\tau(x_1), \_\_\_, S_\tau(x_n)\right)$

$$S_\tau(x) = \begin{cases} x-\tau & x>\tau \\ 0 & |x|\leq\tau \\ x+\tau & x<-\tau \end{cases}$$



Shrinkage

soft-threshold

(#) called the proximal map

$(I+\partial\|\,\|_1)(\underline{x})$

Next, we want to solve:

(#) $A^* = \underset{A}{\operatorname{argmin}} \; \frac{1}{2}\|X-A\|_F^2 + \tau\|A\|_*$

Suppose: $X = U\cdot\Sigma\cdot V^T \rightarrow$ $U^TU=I$, $V^TV=I$

Define: $B = U^TAV \Leftrightarrow A = UBV^T$

(#) is equivalent to:

$B^* = \underset{B}{\operatorname{argmin}} \; \frac{1}{2}\|U(\Sigma-B)V^T\|_F^2 + \tau\|UBV^T\|_*$

CLAIM:

$M \in \mathbb{R}^{m \times n}$, singular values $\sigma_1, \_\_, \sigma_r$

Take $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ - orthogonal

$(U^T U = I, \quad V^T V = I)$

$$\tilde{M} = U M V^T$$

then $\tilde{M}$ has singular values $\sigma_1, \_\_, \sigma_r$

(same).

PROOF

$M = U_M \cdot \Sigma_M \cdot V_M^T$ - SVD of M.

$$\Rightarrow \tilde{M} = \underbrace{(U \cdot U_M)}_{\tilde{U}} \Sigma_M \cdot \underbrace{(V_M^T V^T)}_{\tilde{V}^T}$$

and: $\tilde{U}^T \tilde{U} = (U \cdot U_M)^T (U \cdot U_M)$

$$= U_M^T \cdot \underbrace{(U^T \cdot U)}_{= I} U_M$$

$$= U_M^T \cdot U_M = I_{m \times m}$$

Similarly: $\tilde{V}^T \tilde{V} = I_{n \times n}$

RECALL:

$$\|M\|_* = \sum_{i=1}^{r} \sigma_i$$

$$\|M\|_F^2 = \sum_{i=1}^{r} \sigma_i^2$$

Conclusion:

$$B^* = \arg\min_B \frac{1}{2}\|\cancel{U}(\Sigma - B)\cancel{V^T}\|_F^2 + \tau \|\cancel{U}B\cancel{V^T}\|_*$$

$$\# = \arg\min_B \frac{1}{2}\|\underbrace{\Sigma}_{=\operatorname{diag}(\sigma_1,\ldots,\sigma_r)} - B\|_F^2 + \tau \cdot \|B\|_*$$

$$= \begin{pmatrix} \sigma_1 & 0 & 0 \\ & \ddots \sigma_r & \\ 0 & & \end{pmatrix}$$

Σ - diagonal $\Rightarrow$ B^* - diagonal.

Suppose: $B = \operatorname{diag}(b_1, \ldots, b_r)$

$$\begin{pmatrix} b_1 & 0 & 0 \\ 0 & \ddots \; b_r & \end{pmatrix}$$

$\#$ is equivalent to:

$$\arg\min_{b_1,\ldots,b_r} \frac{1}{2}\sum_i (\sigma_i - b_i)^2 + \tau\left(\sum_{i=1}^r b_i\right)$$

$$\Updownarrow$$

$$\underline{b}^* = \arg\min_{\underline{b}} \frac{1}{2}\|\underline{\sigma} - \underline{b}\|_2^2 + \tau\|\underline{b}\|_1$$

$\underline{b} = (b_1, \ldots, b_r)$ $\quad \underline{\sigma} = (\sigma_1, \ldots, \sigma_r)$

$$\Downarrow$$

$$\underline{b}^* = S_\tau(\underline{\sigma})$$

$X = U\Sigma V^T$

$A = UBU^T$

CONCLUSION:

$$A^* = \arg\min_A \frac{1}{2}\|X - A\|_F^2 + \tau\|A\|_*$$

$$\Rightarrow A^* = U \cdot S_\tau(\Sigma) \cdot V^T = D_\tau(X)$$

singular value threshold

Back to Robust PCA

How to solve:

(i) $L^{(k+1)} = \underset{L}{\operatorname{argmin}}\ \mathcal{L}(L, E^{(k)}, \Lambda^{(k)})$

(ii) $E^{(k+1)} = \underset{E}{\operatorname{argmin}}\ \mathcal{L}(L^{(k+1)}, E, \Lambda^{(k)})$

?

$$\mathcal{L}(L, E, \Lambda) = \|L\|_* + \alpha \|E\|_1$$

$$+ \underbrace{\langle \Lambda, X - L - E \rangle}_{} + \frac{\beta}{2} \|X - L - E\|_F^2$$

(i) Define: $M = X - E^{(k)}$ $\quad \Lambda = \Lambda^{(k)}$

Then:

$$L^{(k+1)} = \underset{L}{\operatorname{argmin}}\ \|L\|_* + \langle \Lambda, M - L \rangle + \frac{\beta}{2} \|M - L\|_F^2$$

$$= \underset{L}{\operatorname{argmin}}\ \frac{1}{2} \|M - L\|_F^2 + \langle \tfrac{1}{\beta} \Lambda, M - L \rangle + \frac{1}{\beta} \|L\|_*$$

$$= \underset{L}{\operatorname{argmin}}\ \frac{1}{2} \|M - L + \tfrac{1}{\beta} \Lambda\|_F^2 + \frac{1}{\beta} \|L\|_*$$

$$\|M - L\|^2 + \langle M - L, \tfrac{1}{\beta} \Lambda \rangle + \frac{1}{\beta^2} \|\Lambda\|^2$$

$$= \underset{L}{\operatorname{argmin}}\ \frac{1}{2} \|\underbrace{M + \tfrac{1}{\beta} \Lambda}_{X} - \underbrace{L}_{A}\|_F^2 + \frac{1}{\beta} \|\underbrace{L}_{A}\|_*$$

singular value thresholding

$$\Rightarrow L^{(k+1)} = D_{\frac{1}{\beta}}\left(M + \tfrac{1}{\beta} \Lambda\right) = D_{\frac{1}{\beta}}\left(X - E^{(k)} + \tfrac{1}{\beta} \Lambda^{(k)}\right)$$

(ii) $E^{(k+1)} = \underset{E}{\operatorname{argmin}} \ \alpha \|E\|_1 + \langle \Delta^{(k)}, X - L^{(k+1)} - E \rangle$
$$+ \frac{\beta}{2} \|X - L^{(k+1)} - E\|_F^2$$

Define: $M = X - L^{(k+1)}$, $\Delta = \Delta^{(k)}$

$$E^{(k+1)} = \underset{E}{\operatorname{argmin}} \ \alpha \|E\|_1 + \langle \Delta, M - E \rangle + \frac{\beta}{2} \|M - E\|_F^2$$

$$= \dots = \underset{E}{\operatorname{argmin}} \ \frac{1}{2} \|(M + \frac{1}{\beta}\Delta) - E\|_F^2 + \frac{\alpha}{\beta} \|E\|_1$$

$$\downarrow$$

$$E^{(k+1)} = S_\tau (M + \frac{1}{\beta}\Delta) = S_\tau (X - L^{(k+1)} + \frac{1}{\beta}\Delta^{(k)})$$

✓ soft threshold

ALGORITM: (ADMM)

Initialisation: $E^{(0)} = \Delta^{(0)} = 0$

Iterate:

- $L^{(k+1)} = D_{\frac{1}{\beta}}(X - E^{(k)} + \frac{1}{\beta}\Delta^{(k)}) \rightarrow$ low rank L
- $E^{(k+1)} = S_{\frac{1}{\beta}}(X - L^{(k+1)} + \frac{1}{\beta}\Delta^{(k)}) \rightarrow$ sparse E
- $\Delta^{(k+1)} = \Delta^{(k)} + \beta(X - L^{(k+1)} - E^{(k+1)})$

Output: $L^{(k)}, E^{(k)}$

$$\underset{x}{\operatorname{argmin}}\ f(x) \quad \text{s.t.} \quad g(x) = 0$$

$$\downarrow$$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda g(x)$$

$$\nabla \mathcal{L}(x, \lambda) = 0$$

Iterative solution:

fixed

(1) $$x^{(k+1)} = \underset{x}{\operatorname{argmin}}\ \mathcal{L}(x, \lambda^{(k)})$$

(2) $$\lambda^{(k+1)} = \lambda^{(k)} + \beta g(x^{(k+1)})$$