21111014

$A \in \mathbb{R}^{n\times d}$ $X \in \mathbb{R}^{d\times p}$ $C \in \mathbb{R}^{n\times p}$ $B \in \mathbb{R}^{q\times p}$.

$\text{Min} \; \|AX - C\|_F$

subject to $\exists\, Y \in \mathbb{R}^{d\times q}$ s.t. $YB = X$

The problem can be stated as

$$\min_Y \|AYB - C\|_F$$

(a). Direct solution

(1) Through physical interpretation

→ $[AYB]$ this matrix has colmn space as a subspace of A's col. sp.

and row space as a subspace of row. sp. of B.

→ To find the matrix $[AYB]$ as close to C as possible it should be equal to the projection of C matrix in col sp. of A and row sp of B

~~Transformation matrix for col sp of A = AA^-~~ ~~Transformation matrix~~

projection matrix onto col. sp. of A = AA^-

projection matrix onto row sp. of B = B^-B

[where $A = U\Sigma V^T$ $\quad B = U'\Sigma' V'^T$

$A^- = V\Sigma^- U^T$ $\quad B^- = V'\Sigma'^- U'^T$

$AA^- = UU^T$ $\quad B^-B = V'V'^T$]

projection of C onto col sp of A and row sp of B.

$\Rightarrow (AA^-)\,C\,(B^-B)$

$\Rightarrow A\,(A^-CB^-)\,B$

So $AYB = A(A^-CB^-)B$ (for Min $\|AYB - C$

$Y = A^-CB^-$

Scanned By Scanner Go

(2) proof of correctness:

consider:

~~AXC~~ $\|AYB - C\|_F^2$ (writing C as $AA^-C + (I - AA^-)C$ in 2 orthogonal spaces)

$$= \|AYB - AA^-C + AA^-C - C\|_F^2$$

$$= \|\underbrace{AA^-}(AYB - C) + \underbrace{(AA^- - I)}C\|_F^2$$

orthogonal spaces

$\left(\begin{array}{l} A^-A = I \\ AYB = AA^-AYB \end{array} \right.$

$$= \|AA^-(AYB - C)\|_F^2 + \|(AA^- - I)C\|_F^2 .$$

$$= \|AYB - \underbrace{AA^-CB^-B - AA^-C(I - B^-B)}\|_F^2 + \|AA^-C - C\|_F^2$$

writing AA^-C in 2 orthogonal spaces B^-B and $I - B^-B$ (~~at~~ orthogonal to each other)

$AYB = AYBB^-B$ since $BB^- = I$.

$$= \|\underbrace{AYBB^-B - AA^-CB^-B} - \underbrace{AA^-C(I - B^-B)}\|_F^2 + \|AA^-C - C\|_F^2$$

$$= \|AYB - AA^-CB^-B\|_F^2 + \|AA^-C(I - B^-B)\|_F^2 + \|AA^-C - C\|_F^2$$

minimised when first term is 0 +ve +ve.

~~AB~~ $AYB = AA^-CB^-B$

$Y = A^-CB^-$

21111014

(b) The embedding function for dimension reduction is given by matrices S (applied from left) and R (applied from right)

$S \Rightarrow k \times n \qquad R \Rightarrow p \times l$.

we have to proove

$$\| SAXBR - SCR \|_F^2 \in (1 \pm O(\epsilon)) \| AXB - C \|_F^2$$

under certain assumption

$$\Rightarrow \| SAXBR - AA^{-}SCR - (I - AA^{-}) SCR \|_F^2$$

$$\Rightarrow \| SAXBR - AA^{-}SCR \|_F^2 + \| (I - AA^{-}) SCR \|_F^2$$

$$+ 2Tr (SAXBR - AA^{-}SCR)^T ((I - AA^{-}) SCR)$$

$$\Rightarrow \| (SAXBR - SAA^{-}CR) + (SAA^{-}CR - SCR) \|_F^2$$

$$\Rightarrow \| SAXBR - SAA^{-}CR \|_F^2 + \| SAA^{-}CR - SCR \|_F^2$$

$$+ 2Tr (SAXBR - SAA^{-}CR)^T (SAA^{-}CR - SCR)$$

$$\Rightarrow \| (SAXBR - SAA^{-}CB^{-}BR) + (SAA^{-}CB^{-}BR - SAA^{-}CR) \|_F^2$$

$$+ \| S(AA^{-}C - C)R \|_F^2 + 2Tr [S [AXB - AA^{-}C] R]^T [S(AA^{-}C - C)R]$$

$$\Rightarrow \| SAXBR - SAA^{-}CB^{-}BR \|_F^2 + \| SAA^{-}CB^{-}BR - SAA^{-}CR \|_F^2$$

$$+ 2Tr (SAXBR - SAA^{-}CB^{-}BR)^T (SAA^{-}CB^{-}BR - SAA^{-}CR)$$

$$+ \| S(AA^{-}C - C)R \|_F^2 + 2Tr (S(AXB - AA^{-}C)R)^T (S(AA^{-}C - C)R)$$

PAGE

DATE

$$\|S(AXB-C)R\|_F^2$$

$$\Rightarrow \|S(AXB - AA^-CB^-B)R\|_F^2 + \|S(AA^-C(I-B^-B))R\|_F^2$$
$$+ \|S(AA^-C-C)R\|_F^2 + 2\,\mathrm{Tr}\left(S(AXB - AA^-CB^-B)R\right)^T \left(S(AA^-C(I-B^-B))R\right)$$
$$+ 2\,\mathrm{Tr}\left(S(AXB - AA^-C)R\right)^T \left(S(AA^-C-C)R\right).$$

~~Considering Term 1~~

and we have $\|AXB-C\|_F^2 = \|AXB - AA^-CB^-B\|_F^2 + \|AA^-C(I-B^-B)\|_F^2 + \|AA^-C-C\|_F^2$.

$\Rightarrow$ now considering term 1

$$\|S(AXB - AA^-CB^-B)R\|_F^2$$
$$= \|SA(X - A^-CB^-)BR\|_F^2$$

Assuming S preserves col. space of A.

i.e.

$$P\left[\forall x \in \mathbb{R}^d, \|SAx\| \in (1 \pm O(\varepsilon))\|Ax\|\right] > 1-\delta$$

for $k = O\left(\frac{d}{\varepsilon^2}\log\frac{1}{\delta}\right)$.

Assumption 1

21/11/2014

We have

$$\| SA(X - A^- C B^-) BR \|_F^2 \in (1 \pm O(\varepsilon))^2 \| A(X - A^- C B^-) BR \|_F^2$$

Assuming R preserves row space of B i.e.

$$P\left[\forall y^T \in \mathbb{R}^q,\ \| y^T B R \| \in (1 \pm O(\varepsilon)) \| y^T B \| \right] \geq 1 - \delta$$

for $l = O\left(\frac{q}{\varepsilon^2} \log \frac{1}{\delta}\right)$

Assumption (2)

We have

$$\| SA(X - A^- C B^-) BR \|_F^2 \in (1 \pm O(\varepsilon))^4 \| A(X - A^- C B^-) B \|_F^2$$

for small values of $O(\varepsilon)$ we have $(1 \pm O(\varepsilon))^4 \approx (1 \pm O(\varepsilon))$

$$\left[\| SA(X - A^- C B^-) BR \|_F^2 \in (1 \pm O(\varepsilon)) \| A(X - A^- C B^-) B \|_F^2 \right]$$

preservation (1)

$\Rightarrow$ now considering term (2)

$$\| S(AA^- C(I - B^- B)) R \|_F^2$$
$$= \| SA(A^- C - A^- C B^- B) R \|_F^2$$
$$\in (1 \pm O(\varepsilon)) \| A(A^- C - A^- C B^- B) R \|_F^2 .$$

Assuming R to be an affine embedding namely

$$P\left[\forall Y \in \mathbb{R},\ \| AA^- C R - YBR \|_F \in (1 \pm O(\varepsilon)) \| AA^- C - YB \|_F \right] > 1 - \delta$$

which requires R to preserve row sp of B and (already assumed)

Assumption (3)

R should also preserve the forbenius norm of AA^-C
i.e.

$$P\left[\|AA^-CR\|_F \in (1 \pm O(\varepsilon))\|AA^-C\|_F\right] > 1-\frac{\delta}{n}$$

if $l = O\left(\frac{1}{\varepsilon^2}\log\frac{n}{\delta}\right)$.

$AA^-C \Rightarrow n \times p$ [preserving norms of n vectors/rows]

Assumption ③

previously

$$\therefore \boxed{l = O\left(\frac{1}{\varepsilon^2}\log\left(\frac{n}{\delta}\right) + \frac{q}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)}.$$

We have under these assumptions :

$$\|SA(A^-C - A^-CB^-B)R\|_F^2 \in (1 \pm O(\varepsilon))^2 \|A(A^-C)R - \underbrace{AA^-CB^-}_{Y}BR\|_F^2$$

$$\in (1 \pm O(\varepsilon))^4 \|AA^-C - \underbrace{AA^-CB^-}B\|_F^2$$

$$\left[\|SA(A^-C - A^-CB^-B)R\|_F^2 \in (1 \pm O(\varepsilon))\|AA^-C - AA^-CB^-B\|_F^2\right]$$

preservation ②

$\Rightarrow$ now consider term ③

$$\|S(AA^-C - C)R\|_F^2$$

Assuming S preserves the norm of $\overbrace{(AA^-C-C)R}^{n \times l}_{n \times l}$ =

$$P\left[S\|(AA^-C-C)R\| \in (1 \pm O(\varepsilon))\|AA^-C-C\|R\right] > 1-\frac{\delta}{l}$$

for $k = O\left(\frac{1}{\varepsilon^2}\log\frac{l}{\delta}\right)$ (since l is small)

$k = O\left(\frac{1}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)$

Assumption ⑤

(already holds) $k = O\left(\frac{d}{\varepsilon^2}\log\left(\frac{1}{\delta}\right)\right)$

21111014

Assuming R preserves the norm of $(\underbrace{AA^{-}C-C}_{n\times p})$

$$P\left[\|(AA^{-}C-C)R\| \in (1 \pm O(\varepsilon))\|AA^{-}C-C\|\right] > 1-\frac{\delta}{n}$$

for $l \geq O\left(\frac{1}{\varepsilon^2}\log\frac{n}{\delta}\right)$

which already holds)

Assumption (4)

we have :

$$\|S(AA^{-}C-C)R\|_F^2 \in (1\pm O(\varepsilon))^2\|(AA^{-}C-C)R\|_F^2$$

$$\in (1\pm O(\varepsilon))^4\|AA^{-}C-C\|_F^2 .$$

$$\|S(AA^{-}C-C)R\|_F^2 \in (1\pm O(\varepsilon))\|AA^{-}C-C\|_F^2$$

preservation (3)

$\Longrightarrow$ now consider term (4)

$$2\,\mathrm{Tr}\left(S(AXB-AA^{-}CB^{-}B)R\right)^T\left(S(AA^{-}C-AA^{-}CB^{-}B)R\right)$$

$$=2\,\mathrm{Tr}\left(S(AXB-AA^{-}CB^{-}B)R\right)\left(S(AA^{-}C-AA^{-}CB^{-}B)R\right)^T$$

$$=2\,\mathrm{Tr}\left(S(AXB-AA^{-}CB^{-}B)R\right)\left(R^T(AA^{-}C-AA^{-}CB^{-}B)^T S^T\right).$$

$$=2\,\mathrm{Tr}\left(SA(X-A^{-}CB)BR\right)\underbrace{\left(R^T(A^{-}C-A^{-}CB^{-}B)^T A^T S^T\right)}_{(SAA^{-}C(I-B^{-}B)R)^T}.$$

$$\Rightarrow 2\,\mathrm{Tr}\left(SA(X-A^{-}CB^{-})\right)\left(BR\,(SAA^{-}C(I-B^{-}B)R)^T\right)$$

$$\leq 2\|SA(X-A^{-}CB)\|_F\,\|BR\,(SAA^{-}C(I-B^{-}B)R)^T\|_F$$

DATE

PAGE

Since $Tr(AB) \leq \|A\|_F \|B\|_F$.

we have:

$$Tr\left(SA(X - A^-CB^-)BR\right)^T \left(SA(A^-C - A^-CB^-B)R\right)$$

$$\leq \underbrace{\|SA(X - A^-CB^-)\|_F}_{\mathrm{I}} \underbrace{\|BR(SAA^-C(I - B^-B)R)^T\|_F}_{\mathrm{II}}$$

Assuming that R preserves the inner product of (q^2 vectors pairs) B_i and $[SAA^-C(I-B^-B)]_{i'}$ (row vectors of respective matrices)

i.e.

$$P\left[\left| B_i (SAA^-C(I-B^-B))^T_{i'} - B_i R (SAA^-C(I-B^-B)_{i'} R)^T \right| \leq O\left(\frac{\epsilon}{\sqrt{q}}\right) \|B_i\| \, \|(SAA^-C(I-B^-B))_{i'}\|\right] \geq 1 - \delta$$

for $i', i = 1, 2, 3 \ldots q/l$ i.e. for q^2 vectors (row) (max)

Assumption 6

II can be written as:

$$\|BR(SAA^-C(I-B^-B)R)^T\|_F^2$$

$$= \|BR(SAA^-C(I-B^-B)R)^T - \underbrace{B(SAA^-C(I-B^-B))^T}_{0} \|_F^2$$

0 as B & $(I - B^-B)$ are orthogonal

$$= \sum_{i=1}^{q} \sum_{i'=1}^{l} \left| B_i R \left((SAA^-C(I-B^-B))_{i'}^T R\right)^T - B_i^T (SAA^-C(I-B^-B))_{i'}^T \right|^2$$

$$\leq \sum_{i=1}^{q} \sum_{i'=1}^{l} \frac{\epsilon^2}{(\sqrt{l})^2} \|B_i\|^2 \|SAA^-C(I-B^-B)_{i'}\|^2$$

21111014

$$= \frac{\varepsilon^2}{q} \|B\|_F^2 \|SAA^-C(I-B^-B)\|_F^2$$

$$= \varepsilon^2 \|SAA^-C(I-B^-B)\|_F^2$$ assuming $B_{q \times p}$ has orthonormal rows

$$\|B\|_F^2 \Rightarrow q.$$

finally we have

$$\|BR(SAA^-C(I-B^-B)R)^T\|_F \leq O(\varepsilon)\|SAA^-C(I-B^-B)\|_F^2$$

$$\in O(\varepsilon)(1 \pm O(\varepsilon))\|AA^-C(I-B^-B)\|_F^2$$

~~(since S preserves col space of A~~

~~& here we are preserving p vectors so~~

$$\cancel{k = O\left(\frac{d}{\varepsilon^2}\log\frac{p}{\delta}\right)}$$

$$\cancel{\therefore k = O\left(\frac{d}{\varepsilon^2}\log\frac{1}{\delta}\right) + O\left(\frac{d}{\varepsilon^2}\log\frac{p}{\delta}\right)}$$

Assuming S preserves the norms of $AA^-C(I-B^-B)$

$$P\left[\|SAA^-C(I-B^-B)\|_F \in (1 \pm O(\varepsilon))\|AA^-C(I-B^-B)\|_F\right] > 1 - \frac{\delta}{p}$$

for $k = O\left(\frac{1}{\varepsilon^2}\log\frac{p}{\delta}\right)$ Assumption (7)

$$\therefore \boxed{k = O\left(\frac{d}{\varepsilon^2}\log\frac{1}{\delta} + \frac{1}{\varepsilon^2}\log\frac{p}{\delta}\right)}$$

$$Tr\left(SA(X-A^-CB^-)BR\right)\left(SA(\cancel{X^*B}\ A^-C - A^-CB^-B)R\right)^T$$

$$\leq \|SA(X-A^-CB^-)\|_F \|BR(SAA^-C(I-B^-B)R)^T\|_F$$

$$\leq (1 \pm O(\varepsilon))^2 O(\varepsilon) \|A(X-A^-CB^-)\| \|AA^-C(I-B^-B)\|$$

$\Longrightarrow$ preservation (4)

$\Rightarrow$ now considering term ⑤

$$Tr\left(S(AXB - AA^{-}C)R\right)^T \left(S(AA^{-}C - C)R\right)$$

$$= Tr\left((XB - A^{-}C)R\right)^T (SA)^T S(AA^{-}C - C)R$$

$Tr(AB) \leq \|A\|_F \times \|B\|_F$

$$\leq \underbrace{\|XBR - A^{-}CR\|_F}_{I} \underbrace{\|(SA)^T S(AA^{-}C - C)R\|_F}_{II}$$

II term is

$$\Rightarrow \|(SA)^T SC^* R - A^T C^* R\|_F^2$$

$C^* = (AA^{-} - I)C$ orthogonal to col sp of A.

$\Downarrow$

0 as $A^T C^* = 0$.

$$\Rightarrow \sum_{j=1}^{d} \sum_{j'=1}^{l} \left|(SA_j)^T (SC^* R_{j'}) - A_j^T (C^* R)_{j'}\right|^2$$

Assuming S preserves the inner product of (d-vectors pairs) A_j and ~~S~~$(C^*R)_{j'}$ (col vectors)

$$P\left[\left|(SA_j)^T (SC^* R)_{j'} - A_j^T C^* R_{j'}\right| \leq O\left(\frac{\epsilon}{\sqrt{d}}\right) \|A_j\| \|C^* R_{j'}\|\right] > 1 - \delta$$

for $j, j' = 1, 2, 3 \ldots d$ & $1, 2, 3 \ldots l$.

Assumption ⑧

$$= \sum_{j=1}^{d} \sum_{j'=1}^{l} \left|(SA_j)^T S(C^* R_{j'}) - A_j^T C^* R_{j'}\right|^2$$

$$\leq \sum_{j=1}^{d} \sum_{j'=1}^{l} \frac{\epsilon^2}{(\sqrt{d})^2} \|A_j\|_2^2 \|C^* R_{j'}\|_2^2$$

Assuming A to have orthonormal cols.

$$= (\epsilon^2/d) \cdot d \, \|C^* R\|_F^2$$

21111014

$$= \varepsilon^2 \| ((AA^- - I)C)R \|_F^2 \in (1 \pm O(\varepsilon))\varepsilon^2 \| (AA^- - I)C \|_F^2$$

under assumption 4

We have :

$$Tr\,(S(AXB - AA^-C)R)^T (S(AA^-C - C)R) .$$

$$\leq \| XBR - A^-CR \|_F \; \| (SA)^T S(AA^-C - C)R \|_F$$

$$\leq \varepsilon^2 (1 \pm O(\varepsilon))^2 \| XB - A^-C \|_F \; \| (AA^- - I)C \|_F^2$$

Since R is an affine embedding given in assumption 3

preservation 5

After considering the assumptions and preservations we have

$$\| SAXBR - SCR \|_F^2$$

$$= \| SA(X - A^-CB^-)BR \|_F^2 + \| SA(A^-CB^-BR - A^-CR) \|_F^2 + \| SAA^-CR - SCR \|_F^2$$

$$+ 2Tr\,(\text{term } 4) + 2Tr\,(\text{term } 5)$$

$$= (1 \pm O(\varepsilon)) \| A(X - A^-CB^-)B \|_F^2 + (1 \pm O(\varepsilon)) \| AA^-C(I - B^-B) \|_F^2 +$$
$$(1 \pm O(\varepsilon)) \| (AA^- - I)C \|_F^2 + (1 \pm O(\varepsilon)) \| A(X - A^-CB^-) \|_F^2 \| AA^-C(I - B^-B) \|_F$$

$$+ (1 \pm O(\varepsilon)) \| AXB - AA^-C \|_F^2 \| (AA^- - I)C \|_F^2$$

24/11/014

Simplifying for small ε we finally arrive at

$$\left| \|SAXBR - SCR\|_F^2 - \|AXB - C\|_F^2 \right| \leq O(\varepsilon) \|AXB - C\|_F^2$$

where $\underset{k\times n}{S_{ij}} \sim N\left(0, \frac{1}{k}\right)$ and $\underset{p\times l}{R_{ij}} \sim N\left(0, \frac{1}{l}\right)$

and $k = O\left(\frac{d}{\varepsilon^2} \log \frac{1}{\delta} + \frac{1}{\varepsilon^2} \log \frac{p}{\delta}\right)$

and $l = O\left(\frac{1}{\varepsilon^2} \log\left(\frac{n}{\delta}\right) + \frac{q}{\varepsilon^2} \log \frac{1}{\delta}\right)$.

optimal solution for $\min_X \|AXB - C\|_F^2$

$$X^* = \underset{(n\times d)^+}{A^-} \underset{(n\times p)}{C} \underset{(q\times p)^+}{B^-}$$

Time complexity $O(n^2 p^2 d q)$

optimal solution in embedded space $\exists$

$$\cancel{\hat{X} = \underset{(k\times d)}{(SA)^-} \underset{k\times l}{SCR} (SB)^-}$$

$$\hat{X} = \underset{(k\times d)^-}{(SA)^-} \underset{(k\times l)}{(SCR)} \underset{(q\times l)^-}{(BR)^-}$$

Time complexity $= O(dq\, k^2 l^2)$.

DATE
PAGE