Short Papers

A-Optimal Projection for Image Representation

, *Senior Member, IEEE*, ,
, *Member, IEEE*,
, *Fellow, IEEE*

Abstract—

2.2 Laplacian Regularized Least Squares

$$y = \mathbf{w}^T\mathbf{x} + \epsilon, \quad (2)$$

y *observation,* $\mathbf{x} \in$ ^{n} *independent variable,* $\mathbf{w}$ *weight vector,* ϵ . D

σ^2. $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x}$,

$\mathbf{x}$ $\mathbf{w}$.

$\{\mathbf{x}_i, y_i\}_{i=1}^{\ell}$, ,

$\widehat{\mathbf{w}}$,

$$J_{\text{sse}}(\mathbf{w}) = \sum_{i=1}^{\ell}\left(\mathbf{w}^T \; {}_i - y_i\right)^2. \quad (3)$$

$$\widehat{\mathbf{w}} = (XX^T)^{-1}X\mathbf{y}, \quad (4)$$

$X = (\mathbf{x}_1, \ldots, \mathbf{x}_\ell)$ $\mathbf{y} = (y_1, \ldots, y_\ell)^T$.

S

. $\mathcal{N}_k(\mathbf{x})$

k $\mathbf{x}$.

S. A :

$$S_{ij} = \begin{cases} 1, & \text{if } \mathbf{x}_i \in \mathcal{N}_k(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in \mathcal{N}_k(\mathbf{x}_i); \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

$$\min_{\mathbf{w}} \sum_{i=1}^{m}(\mathbf{w}^T\mathbf{x}_i - y_i)^2 + \frac{\lambda_1}{2}\sum_{i,j=1}^{m}(\mathbf{w}^T\mathbf{x}_i - \mathbf{w}^T\mathbf{x}_j)^2 S_{ij} + \lambda_2\|\mathbf{w}\|^2,$$

$$\widehat{\mathbf{w}} = (XX^T + \lambda_1 XLX^T + \lambda_2 I)^{-1}X\mathbf{y}, \quad (6)$$

$L = diag(S\mathbf{1}) - S$ *graph Laplacian* 23 **1**

2.3 Optimal Experimental Design

C (2) (4). B

, $\hat{\mathbf{w}} - \mathbf{w}$

$\sigma^2 H_{\text{sse}}^{-1}$, H_{sse}

J_{sse} (3):

$$H_{\text{sse}} = \left(\frac{\partial^2 J_{\text{sse}}}{\partial \mathbf{w}^2}\right) = \left(\sum_{i=1}^{\ell} \; {}_i \; {}_i^T\right) = XX^T.$$

Optimal Experimental Design 16 *optimal*

size

$\hat{\mathbf{w}} - \mathbf{w}$.

:

- D- : ;
- A- : ;
- \- : .

3 A-OPTIMAL PROJECTION

, A-

.

3.1 Problem Definition

$X = (\mathbf{x}_1, \ldots$

(11) :

$$\min_A \mathrm{Tr}\big((A^T\widetilde{X}\widetilde{X}^TA+\lambda_2 I)^{-1}\big). \tag{12}$$

:

3.1. *The optimization problem (12) is equivalent to the following:*

$$\min_A \mathrm{Tr}\big((\widetilde{X}^TAA^T\widetilde{X}+\lambda_2 I)^{-1}\big). \tag{13}$$

C

D ://. . /10.1109/ A .2015.2439252.

(13),

$\|A\|_F \le \rho$ $\gamma\|A\|_F^2$

3.3 Relations to Other Methods

A

L, S.

i j

24 .

z.

B

2.1,

. A *principal component regression* (C) 25 . *principal component analysis* (CA)

λ_1 λ_2 (11),

$$\min_A \mathrm{Tr}((A^TXX^TA)^{-1}). \tag{14}$$

CA.

C

. *Partial least squares regression* ( ) 26

x z

C

, *sliced inverse regression* ( ) 27

x z

x. $z \perp x|y$, y

x.

x z.

x,

28 .

, $v_1 \geq \cdots \geq v_n \geq 0$. ,
A

$$A = (\mathbf{a}_1, \ldots, \mathbf{a}_k), \ \mathbf{a}_i = \sqrt{v_i}\mathbf{u}_i. \tag{19}$$

4.2 Iterative Optimization

A

, D

.

,

.

:

4.3. *The optimization problem (13) is equivalent to the following optimization problem:*

$$\min_{A,B} \|I - A^T \widetilde{X} B\|^2 + \lambda \|B\|^2, \tag{20}$$

where $A \in \mathbb{R}^{n \times k}$ and $B \in \mathbb{R}^{m \times k}$ are the two variables.

4.3 A - A B. B . D - (20) ϕ. B . B ϕ B^T ,

$$\begin{aligned} &\frac{\partial \phi}{\partial B^T} = 0 \\ &\Rightarrow B^T \widetilde{X}^T A A^T \widetilde{X} + B^T - A^T \widetilde{X} = 0 \\ &\Rightarrow B = (\widetilde{X}^T A A^T \widetilde{X} + \lambda I)^{-1} \widetilde{X}^T A. \end{aligned} \tag{21}$$

B . A, ϕ A :

$$\begin{aligned} &\frac{\partial \phi}{\partial A} = 0 \\ &\Rightarrow -2 \frac{\partial \mathrm{Tr}(A^T \widetilde{X} B)}{\partial A} + \frac{\partial \mathrm{Tr}(A^T \widetilde{X} B B^T \widetilde{X}^T A)}{\partial A} = 0 \\ &\Rightarrow -2\widetilde{X} B + 2\widetilde{X} B B^T \widetilde{X}^T A = 0 \\ &\Rightarrow A = (\widetilde{X} B B^T \widetilde{X})^{-1} \widetilde{X} B. \end{aligned} \tag{22}$$

A

:

1) A CA , , ;
2) C B . (21);
3) A $\|A\|_F \leq \rho$
   - A . (22),
   - A ;

else if we control the size of A via a regularizer $\gamma \|A\|_F^2$
   - A

$$A = (\widetilde{X} B B^T \widetilde{X} + \gamma I)^{-1} \widetilde{X} B.$$

4) 2 3 .

5 Experimental Results

, .

.

5.1 Relevance Feedback Image Retrieval

\-

\- (CB) 30 . , -

\- :

- (, 21),
  2.1.
- A (A , 10).
  21 , A
  , .
- (, 22). A ,
  .
- A . ,
  :

$$S_{ij}^{AOP} = \begin{cases} \alpha, & \text{if } \mathbf{x}_i \in \boldsymbol{F}^+ \text{ and } \mathbf{x}_j \in \boldsymbol{F}^+; \\ 1, & \text{if } \mathbf{x}_i \in \mathcal{N}_k(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in \mathcal{N}_k(\mathbf{x}_i); \\ 0, & \text{otherwise}, \end{cases} \tag{23}$$

α .

A . ,
4.2,
A.

5.2 Data Preparation

5,000 50 -
C .
. , 64-
64- (C) 31 -
. $4 \times 4 \times 4$
. C , . 31 ,
. C
\-
(, ,) .
, C
.
31 .

5.3 Evaluation Settings

precision-scope curve *precision rate* 32
. -
N - .
\-
N. -

.
,
. , 20
, 20
.
five fold cross validation -
. ,
. , 20
. A , -
,
. -

\- .

. ,
. 10
,

(... %)

C	B	A	A			C	B	A	A		
	99.3	**99.8**	99.8	99.8	99.8		38.2	**70.5**	61.7	52.0	66.2
C	94.0	99.3	98.3	**99.8**	99.3	B	38.5	**58.2**	45.8	35.0	43.0
D	87.8	98.7	98.5	97.2	**99.3**	C	39.0	**60.3**	47.5	27.2	48.2
	85.8	95.8	90.8	**96.7**	91.0	D	37.5	**50.0**	45.5	32.0	32.0
	83.3	**95.7**	70.7	88.7	66.7	C	35.3	**55.5**	49.8	36.5	47.3
C	73.0	**94.8**	93.8	89.5	94.2		40.0	**58.2**	46.5	40.5	42.7
D	67.5	**92.5**	79.5	90.0	87.5		33.5	**56.5**	54.8	44.3	52.8
	74.0	**95.0**	88.5	90.8	89.8		51.0	**71.5**	48.7	44.0	45.8
	79.2	**80.3**	80.3	80.3	80.3		32.2	**49.5**	43.3	43.2	39.8
	55.5	82.0	78.5	70.0	**82.0**		32.5	**47.0**	45.8	25.5	45.8
	53.3	**82.8**	59.7	58.5	67.5	B	28.5	**55.0**	50.3	33.5	46.3
	51.7	**74.8**	53.0	45.7	48.0	**46.8**	24.8	45.0	**46.8**	31.3	38.8
	43.5	**74.8**	61.0	57.0	66.8		32.5	**52.5**	45.8	40.5	42.8
C	62.0	**78.8**	63.3	63.0	63.3	B	21.5	**42.3**	36.0	18.5	36.5
A	44.3	**66.8**	56.3	48.5	51.2		39.0	**57.0**	54.5	35.37.9699153.00713	

( ) - . - - 15 . , , +1 −1, . . , . . , . . , 300 , , - . λ_1 λ_2 10^{-4}. λ_1 λ_2 A . - , 2 50 .

5.4 Image Retrieval Performance

, . . 1

20 - 50 . *baseline* . . , 128- - , . A , A , A , , - . A , 50 , A 43 . A , A , 65.3, 57.7, 56.2, 49.4 , - . C , , A , A 7.6 . , , 50 , - . , , , - . A , , . 1 *precision-scope* . A , A - . A , . A , - . A ,

. A . A A , A . A A , A ,

20, 30, 50) , (10,

. 2. A

20, 45.7 90.1

. A 45.7 80.7 81.3

5.5 Parameter Selection

λ_1 λ_2.

20 . 3 3 .

. 3 , $\lambda_2 = 10^{-4}$ λ_1

; . 3 , $\lambda_1 = 10^{-4}$

λ_2 . A

6 Conclusions

A-

Acknowledgments

B C (973) 2012CB316400, C 61233011 61125203.

References

1 *IEEE Trans. Pattern Anal. Mach. Intell.*, . 27, . 3, . 328 340, . 2005.

2 D. *Proc. IEEE Int. Conf. Data Min.*, , 2008, . 433 442.

3 D. *IEEE Trans. Pattern Anal. Mach. Intell.*, . 31, . 2, . 260 274, . 2009.

4 . C C *Proc. 5th ACM SIGMM Int. Workshop Multimedia Information Retrieval*, B , CA, A, 2003, . 193 200.

21 . . , , *Proc. Adv. Neural Inf. Process. Syst. 16*, C , A, A, 2003, . 153 160.

22 . . , - , *Proc. 14th Annu. Int. ACM Conf. Multimedia*, B , CA, A, ., 2006, . 297 306.

23 . . . C , *Spectral Graph Theory* (C), , , A: A , . 92, 1997.

24 D. C , . , . , - , *Proc. IEEE 11th Int. Conf. Comput. Vis.*, 2007, . 1 8.

25 . . ff , A , *J. Roy. Statist. Soc.*, . 31, . 3, . 300 303, 1982.

26 . B. . . (1986). - : A . *Analytica Chimica Acta* . *185*, . 1 17. A : :// . . / / / /0003267086800289

27 .-C. . (1991). . *J. Am. Statist. Assoc.* . *86(414)*, . 316 327. A : :// . . / /2290563

28 . , . . B , . . , , *Ann. Statist.*, . 37, . 4, . 1871 1905, 2009.

29 . , C. C , . C , . B , D. C , . , C - , *Proc. 17th ACM Int. Conf. Multimedia*, 2009, . 45 53.

30 A. . , . , . , A. , . , C - , *IEEE Trans. Pattern Anal. Mach. Intell.*, . 22, . 12, . 1349 1380, D . 2000.

31 . , . , .- . , . , C - , *Proc. Int. Conf. Image Process.*, 2002, . 24 28.

32 D. . . , - : A , *IEEE Trans. Pattern Anal. Mach. Intell.*, . 27, . 2, . 245 251, . 2005.