2014 IEEE International Conference on Systems, Man, and Cybernetics
October 5-8, 2014, San Diego, CA, USA

DYNAMICALLY REGULARIZED HARMONY LEARNING OF GAUSSIAN MIXTURES

Hongyan Wang, Jinwen Ma

De a me f I f ma i Scie ce, Sc l f Ma ema ical Scie ce
e i g U i e i , ei i g, 100871, C i a

ABSTRACT

I i a e, a d amicall eg la i ed a m lea i g D alg i m i ed f a ia mi e lea i g i a fa i e fea e f b ada i e m del elec i a d c i e a ame e e ima i . S eci call , de e f ame f a e ia i g- a g a m lea i g, e ili e e a e age S a e f e e i babili e am le a a eg la i a i e m be i g c lled b a cale fac e a m f c i a ia mi e i c ea i g f m 0 1 d amicall . I i dem a ed b e e ime b e ic a d eal- ld da a e a e D alg i m ca l elec e c ec mbe f ac al a ia i e da a e b al b ai e ma im m li eli d M e ima f e a ame e i e ac al mi e. M e e , e D alg i m i calable a d ca be im leme ed a big da a e .

Keywords a ia mi e m del elec i eg la i a i ma im m li eli d.

1. INTRODUCTION

A a e ible a d e f l a i ical l f da a a al i a d i f ma i ce i g, i e mi e m del 1 a f d i a lica i i ma blem , c a cl e i g a al i , image egme a i a d eec ec g i i . A m g e e a lica i , a ia mi e a e idel ed a d e e al a i ical lea i g me d a e bee ed deal i i i d f m del , c a e EM alg i m 2 a d e me d f m me 3 . e all a med a e mbe f a ia cl e i a da a e i e- . e e, i ma i a ce i e i f ma i i a ailable. e , e elec i f a i a e mbe f a ia , called m del elec i , m be made i e a ame e lea i g i mi e . , e ge e al a ia mi e m deli g i ac all a c m d m deli g blem f b a ame e e ima i a d m del elec i , ic i a a e c m lica ed a d dif c l a 4 .

e c e i al a f l i g i c m d mi e m deli g blem i elec a imal mbe k^* f a ia a e cl e i e da a e ia e f e i f ma i , c di g a d a i ical elec i c i e ia c a A ai e I f ma i C i e i 5 , a e ia I fe e ce C i e i IC , Mi im m De c i i e g MD 7 , a d Mi im m Me age e g MM 8 . e e, e alida i g ce f i me d i c m a i all e e i e beca e eed e ea e e i e a ame e lea i g ce a a la ge mbe f ible k.

Si ce e 1990 , me a i ical lea i g a ac e a ea ed l e i blem. Di ic le ce e 9 a d e e ible m Ma c ai M e Ca l MCMC 10 a e ical im leme a i f e i d fa ac e ic e c a ic im la i . e e c a ic im la i me d ge e all elie i e i e am li g a d a e e ime-c mi g. e ec d i di e a e ia m del ea c ba ed imi i g e a ia i al b d 11 . e i d i e i ed lea i g i e mi e 12, 13 ic i d ce ce ai c m e i i e mec a i m i e mi e m del c a e m del elec i ca be made ada i el d i g a ame e lea i g i a im li ed MM m del elec i c i e i i g e C mec a i m.

Al e a i el , e a e ia i g- a g a m lea i g 14, 15, 1 a al ided a e a i ical lea i g mec a i m a ma e m del elec i ada i el d i g a ame e lea i g. I a al ead bee im leme ed a ia mi e lea i g a d e e al a m lea i g alg i m a e al bee e abli ed f a ia mi e 17, 18, 19 . Al g e a m lea i g e abili f ada i e m del elc i , i a ame e e ima i a a able de ia i f m e M e ima i ic i c i e i e a ame e .

I de l e i de ia i blem, e c ide e a m lea i g a e M lea i g i a eg la i a i e m f a e age ega i e S a e f e e i babili e am le i g c l e cale c m le i f mi e m del . m i i f ie , e e a d amicall eg la i ed a m lea i g D alg i m f a ia mi e b addi g a S a e eg la i a i e m bei g c lled b a cale fac e a m lea i g. A e cale fac i i c ea ed d amicall f m 0 1, ed al-

i e Ma, e c e di g a , el 8 -10- 27 0 09, Email ma ma . .ed .c .

g i m a f m f m e a m lea i g i ada i e m del elec i i e c e i al ma im m li eli d lea i g a e ada i e m del elec i a d e M e ima i a e b b ai ed a la .

e e f e a e i ga i ed a f ll . e begi i a b ief de c i i f e a m lea i g em f a ia mi e i Sec i 2. e, e e e e de i a i a d a al i f e d amicall egla i ed a m lea i g alg i m f a ia mi e i Sec i 3. Sec i 4 c ai e e e ime al e l b e ic a d eal- ld da a e . i all , e c cl de b ie i Sec i 5.

2. BYY HARMONY LEARNING OF GAUSSIAN MIXTURES

e a m lea i g em de c ibe eac b e-a i $x \in X \subset \Re^n$ a d i c e di g i e e e-e a i $y \in Y \subset \Re^m$ ia e e f a e ia dec m i i f e i de i $p(x,y) = p(x)p(y|x)$ a d $q(x,y) = q(y)q(x|y)$, ic a e called a g mac i e a d i g mac i e, e ec i el . i e a am le da a e $D_x = \{x_t\}_{t=1}^N$ f m e a g b e able ace, e a m lea i g em i i g e ac e id-de babili ic c e f x i e el f y f m ec-if i g all a ec f $p(y|x)$, $p(x)$, $q(x|y)$ a d $q(y)$ b ma i-mi i g e f ll i g a m f c i al

$$H(p||q) = \int p(y|x)p(x)\ln[q(x|y)q(y)]dxdy. \quad (1)$$

If b $p(y|x)$ a d $q(x|y)$ a e a ame ic, e lea i g em i called a e a i-di ec i al A c i ec e i-A c i ec e f . i e a am le da a e $D_x = \{x_t\}_{t=1}^N$, e i-a c i ec e f e a m lea i g em ca be eci ed a f ll . e i e e e e a i y i di c e e i $Y = \{1, 2, \cdots, k\}$ i.e., i $m = 1$, ile e b e a i x i c i f m a a ia mi e di ib i . O e i g ace, e le $q(y = j) = \pi_j \geq 0$ i $\sum_{j=1}^k \pi_j = 1$. i i a i babili di ib i f a ia cl e f e mi e. O e a g ace, $p(x)$ i a la e babili de i f c i df f a ia mi e f m ic D_x a e ge e a ed. M e e, i e i g a , $q(x|y = j) = q(x|m_j, \Sigma_j)$ i a med be a a ia de i f c i i mea ec m_j a d e c a ia ce ma i Σ_j, ile i e a g a , $p(y = j|x)$ i c c ed de e a e ia i ci le b e f ll i g a ame ic f m,

$$p(y = j|x) = \frac{\pi_j q(x|m_j, \Sigma_j)}{q(x|\Theta_k)}, \quad (2)$$

$$q(x|\Theta_k) = \sum_{j=1}^k \pi_j q(x|m_j, \Sigma_j), \quad (3)$$

e e $\Theta_k = \{\pi_j, m_j, \Sigma_j\}_{j=1}^k$ a d $q(x|\Theta_k)$ i a a-ia mi e m del a ill a ima e e la e $p(x)$ ia e a m lea i g e lea i g em.

all e e c m e i E . 1 , e a e

$$H(p||q) = E_{p(x)}\left[\sum_{j=1}^k h_j(X)\ln[\pi_j q(X|m_j, \Sigma_j)]\right], \quad (4)$$

e e

$$h_j(X) = \frac{\pi_j q(X|m_j, \Sigma_j)}{\sum_{i=1}^k \pi_i q(X|m_i, \Sigma_j)}. \quad (5)$$

a i , $H(p||q)$ i e e ec a i f a f c i f e a-d m a iable X b ec $p(x)$. i e am le da a e D_x, e ge a e ima e f $H(p||q)$, called a m f c-i , a f ll

$$J(\Theta_k) = \frac{1}{N}\sum_{t=1}^N\sum_{j=1}^k h_j(x_t)\ln[\pi_j q(x_t|m_j, \Sigma_j)].$$

Acc di g e e e ical a d e e ime al e l i i-a c i ec e f e a m lea i g em f a ia mi e 20, 17, 18, 19 , e ma imi a i f $J(\Theta_k)$ i ca able f ma i g m del elec i ada i el d-i g a ame e lea i g e e ac al a ia cl-e a e e a a ed i a ce ai deg ee. a i , if e c e k be la ge a e mbe k^* f ac al a ia cl e i e am le da a, e ma imi a i f e a m-f c i ca ma e k^* a ia ma c e ac al e a d im l a e l elimi a e $k - k^*$ e a e . e e , a e me i ed e i l , e igi al a m lea i g ffe f m i c i e a ame e e ima i . S , e e i e e eg la i a i mec a i m a -f m e a m lea i g e M lea i g c a ada i e m del elec i a d c i e a ame e e i-ma i ca be made im l a e l .

3. DYNAMICALLY REGULARIZED HARMONY LEARNING ALGORITHM

3.1. The Dynamic Regularization Mechanism

Acc di g 21 , $J(\Theta_k)$ ca be di ided i a ,

$$J(\Theta_k) = L(\Theta_k) - O_N(p(y|x)), \quad (7)$$

e e e a i e l g-li eli d f c i , i.e.,

$$L(\Theta_k) = \frac{1}{N}\sum_{t=1}^N \ln(\sum_{j=1}^k(\pi_j q(x_t|m_j, \Sigma_j))), \quad (8)$$

ile e ec d i e a e age S a e f e -e i babili $p(y|x)$ e e am le da a e $\mathcal{D} = \{x_t\}_{t=1}^N$,

$$O_N(p(y|x)) = -\frac{1}{N}\sum_{t=1}^N\sum_{j=1}^k p(j|x_t)\ln p(j|x_t). \quad (9)$$

Acc di g E . 7 , if $-O_N(p(y|x))$ i ie ed a a eg- la i a i e m, e a m lea i g, i.e., ma imi - i g $J(\Theta_k)$, i a eg la i ed M lea i g ic a al ead bee i e iga ed i 22, 23, b cali g e eg la i a i e m i a mall i i e mbe . e e, i ce e ee e eg la i a i cale c a a i e ca e f e a m lea i g, e e i e iga i al ffe f m i c i e a ame e e ima i .

O e e a d, f m E . 7 e al a e

$$L(\Theta_k) = J(\Theta_k) + O_N(p(y|x)), \qquad 10$$

ic i dca e a e M lea i g i a eg la i ed a m lea i g i $O_N(p(y|x))$ a e eg la i a i e m. c l e eg la i a i , a cale fac $\lambda(\geq 0)$ i i d ced,

$$L_\lambda(\Theta_k) = J(\Theta_k) + \lambda O_N(p(y|x)). \qquad 11$$

If $\lambda = 0$, $L_\lambda(\Theta_k) = J(\Theta_k)$ i a m f c- i e i-a c i ec e f a ia mi e . If $\lambda = 1$, $L_\lambda(\Theta_k)$ i e l g-li eli d f c i f e a ia mi - e m del. a i , i λ i c ea i g f m 0 1, ma imi - i g $L_\lambda(\Theta_k)$ c a ge f m e a m lea i g e M lea i g. e e e c l e i c ea i g f λ a - ia el eali e ada i e m del elec i a e e i lea i g age a d e M e ima i a e al lea i g age.

3.2. The Fixed-point Learning Algorithm

A eac a e f ed amicall eg la i ed a m lea - i g i a eci c λ, e c c a ed- i alg i m ma imi e $L_\lambda(\Theta_k)$.

c e ie ce, e ili e e f ma e e e a i f π_j, i.e., $\pi_j = e^{\beta_j}/\sum_{i=1}^{K} e^{\beta_i}$, $j = 1, \cdots, k$, e e $\beta_j \in (-\infty, +\infty)$, $j = 1, \cdots, k$. e i g e de i a i e f $L_\lambda(\Theta_k)$ i e ec β_j, m_j a d Σ_j, e ec i el , be e , e ge e f ll i g ed- i i e a i e lea i g alg i m

$$\hat{\pi}_j = \frac{\sum_{t=1}^{N} p(j|x_t)\gamma_j(t)}{\sum_{t=1}^{N}\sum_{i=1}^{k} p(i|x_t)\gamma_i(t)}; \qquad 12$$

$$\hat{m}_j = \frac{\sum_{t=1}^{N} p(j|x_t)\gamma_j(t)x_t}{\sum_{t=1}^{N} p(j|x_t)\gamma_j(t)}; \qquad 13$$

$$\hat{\Sigma}_j = \frac{\sum_{t=1}^{N} p(j|x_t)\gamma_j(t)(x_t - \hat{m}_j)(x_t - \hat{m}_j)^T}{\sum_{t=1}^{N} p(j|x_t)\gamma_j(t)}, \qquad 14$$

e e

$$\gamma_i(t) = 1 - \sum_{l=1}^{k}(p(l|x_t) - \delta_{il})\ln \pi_l p(x_t|m_l, \Sigma_l) + \lambda \sum_{l=1}^{k}(p(l|x_t) - \delta_{il})\ln p(l|x_t), \qquad 15$$

e e δ_{ij} i e ec e f c i .

I c m a i i e c e i al EM alg i m f a ia mi e 2 , ed ed- i lea i g al- g i m diffe l a e a gme i g e m $\gamma_j(t)$. I ca be ea il e i ed a e $\lambda = 1$, $\gamma_j(t) = 1$, e ed- i lea i g alg i m i e EM alg i m a d e $\lambda = 0$, e ed- i lea i g alg i m e e - igi al ed- i lea i g alg i m f ma imi i g e a m f c i $J(\Theta_k)$.

Ac all , $\gamma_j(t)$ im leme a i al e ali ed c m e i- i e lea i g C mec a i m 24 - 25 a m del - elec i ca be made ada i el d i g a ame e lea i g. A e ea l lea i g age, $\gamma_j(t) < 0$ ma a e . Acc d- i g E . 15 , e mea ec f j- a ia ill m e a a f m x_t. O e i e, if $\gamma_j(t) > 0$, e mea ec f e j- a ia ill be a ac ed x_t. S , f x_t, a - ia i $\gamma_j(t) > 0$ a e i e ile e e a ia i $\gamma_j(t) < 0$ a e l e .

e e, e ed- i lea i g alg i m ca g a - a ee e i i e de i e e f eac c a ia ce ma i d i g e i e a i i ce $\gamma_j(t)$ ma be ega i e. I de e c me i blem, e e e EM da e le f e c a ia ce ma i e , i.e., f ci g all $\gamma_j(t) = 1$ i E . 14 , i i eci c ca e. I fac i im li ca i i a lica- ble a d ef cie i ce e c m e i i f ada i e m del elec i i mai l am g mea ec a d c lled b e mi i g i .

3.3. The Dynamic Evolution of λ

e f e di c ed amic e l i f λ i ime T i g e lea i g ce . Acc di g eg la i a i mec a i m, λ ld be e mall a d i c ea e l l a e ea l lea i g age eali e ada i e m del elec i . e , a e e e age, λ ca g l a a ic e eed a d e alg i m ill all c e ge a M l i . S , i i c cial c ec e e e ada i e m del elec- i a acc m li ed a d e c a ge lea i g age.

I de de ec e i g i e i d ce e S a e f mi i g i i a ia mi - e , $H_\pi = -\sum_{j=1}^{k}\pi_j \ln \pi_j$. Ob i l , H_π i e i i e e c e f a mi e m del. If m del elec i i c m le ed, e diffe e ce f H_π be ee i e a i i c ide able. O e i e, e diffe e ce ld be e mall. i m i a e ad e ab l e c a ge a e f H_π be ee i e a i , de ed b

$$h_\pi(T) = \left|\frac{H_\pi(T) - H_\pi(T-1)}{H_\pi(T)}\right|, \qquad 1$$

a a i dica f m del elec i . T i ime, i.e., e m- be f i e a i e le lea i g ce i di ided i lea i g age acc di g a gi e e ld $\varepsilon_1(> 0)$ f i i dica . a i , if $h_\pi(T) > \varepsilon_1$, $\lambda(T)$ i c ea e a l w eed, e wi e, i i c ea e a ig eed. Si ce $\lambda(T)$ i a med i c ea e e e iall , i d amic e l i ce i gi e a f ll w,

$$\lambda(T) = \begin{cases} \lambda_0 * \eta_1^T, & \text{if } h_\pi(T) > \varepsilon_1; \\ \lambda_0 * (\frac{\eta_1}{\eta_2})^{T^*} \eta_2^T, & \text{if } h_\pi(T) \leq \varepsilon_1, \end{cases} \quad 17$$

w e e λ_0 bei g a e mall i i e c a i i ial al e f λ, η_1, η_2 a e w i i e c a wi c ai a $1 < \eta_1 < \eta_2$, a d T^* i e i g i c a $h_\pi(T^*) > h_0$ a d $h_\pi(T^* + 1) \leq h_0$. 4 e λ eac e 1, we i il e i e a i c e ge .

3.4. The Complete DRHL Algorithm

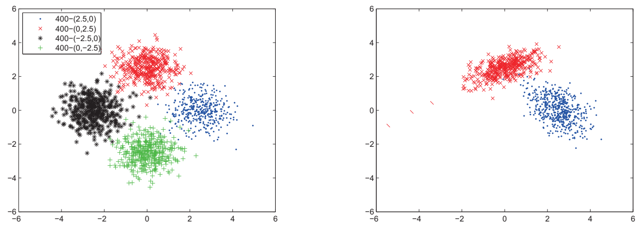
4e all mma ed D alg i m. i l , we ld c e e a ame e f e alg i m e l . A me i ed e i l , $\lambda_0, \eta_1, \eta_2$ a d ε_1 m be ca e- f ll elec ed ma e e e l i f $\lambda(T)$ d amic. θ_0 i a e ld al e l e a ia wi e mal- l mi i g i d i g e a ame e lea i g ce , w ile $\varepsilon_2(> 0)$ i a e ld al e e mi a e e i e a i . If $\lambda = 1$ a d e ab l e i c eme f e l g li eli d i malle a ε_2, we af m e c e ge ce f e alg - i m. I lea i g a adigm, k i e ible. w e e, i ld be la ge a e mbe k^* f ac al a ia - cl e i e da a e . A f e i i ial e i g f e a ame e Θ_k, i.e., $\Theta_k^{(0)} = \{\pi_i^0, m_i^0, \Sigma_i^0\}|_{i=1}^k$, me adi- i al cl e i g me d ma be el f l. e am le, m_i^0 ca be elec ed g a C ced e 25 a d e π_i^0 a d Σ_i^0 ca be e ima ed acc di gl .

Af e i i iali i g all e a ame e , Θ_k will be da ed i eac a e f $\lambda(T)$ ia e ed- i lea i g alg i m gi e b E 12 - 14 . A e e d f eac lea i g a e, e a ia wi e mi i g i le a λ_0 a e a i ila ed immedia el . Af e $\lambda(T)$ eac e 1, e alg - i m g e il e l g li eli d f c i eac e i ma im m al e i ab l e i c eme i le a ε_2.

Si ce e D alg i m a eac lea i g a e be- c me e ed- i lea i g alg i m w ic a e im- ila da e le a e EM alg i m, we ca e e da a mma i a i ec i e gge ed f e EM alg i m f a ia mi e i 2 , ma e i calable a d be im- leme ed a big da a e e ef e, e D alg i m ca be calable a d ed a big da a e

4. EXPERIMENTAL RESULTS

I i ec i , a i e e ime a e ca ied b e ic a d eal-w ld da a e dem a e e e- f ma ce f e d amicall eg la i ed a m lea i g D alg i m f a ia mi e . M e e, i i c m a ed wi me ical e i i g lea i g alg i m- . I e e e e ime , we al wa elec $\varepsilon_1 = 1e-5$,



a

[redacted]

[redacted]

Table 1. e al e f e a ame e f e f e ic da a e .

e da a e	a ia	m_i	σ^i_{11}	$\sigma^i_{12}(\sigma^i_{21})$	σ^i_{22}	π_i	N_i
$\mathcal{S}_1$	1	2.50,0	0.50	0.00	0.50	0.25	400
(−1 00)	2	0,2.50	0.50	0.00	0.50	0.25	400
	3	-2.50,0	0.50	0.00	0.50	0.25	400
	4	0,-2.50	0.50	0.00	0.50	0.25	400
$\mathcal{S}_2$	1	2.50,0	0.45	-0.25	0.55	0.34	544
(−1 00)	2	0,2.50	0. 5	0.20	0.25	0.28	448
	3	-2.50,0	1.00	0.10	0.35	0.22	352
	4	0,-2.50	0.30	0.15	0.80	0.1	2 5
$\mathcal{S}_3$	1	2.50,0	0.10	-0.20	1.25	0.50	00
(−1200)	2	0,2.50	1.25	0.35	0.15	0.30	3 0
	3	-1,-1	1.00	-0.80	0.75	0.20	240
$\mathcal{S}_4$	1	2.50,0	0.28	-0.20	0.32	0.34	8
(−200)	2	0,2.50	0.34	0.20	0.22	0.28	5
	3	-2.50,0	0.50	0.04	0.12	0.22	44
	4	0,-2.50	0.10	0.05	0.50	0.1	32

g e ell a k^* a ia dem a ed b ei c li e a e all ec g i ed a d eac e ima ed a ia ma c e e ac al e acc a el .

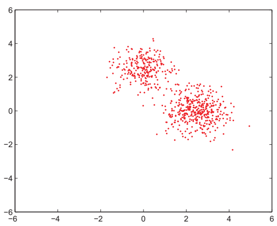


Table 2. e c m a i f e D a d CEM^2 alg i m m del elec i a d ime.

Da a e	D		CEM^2	
	CMS e e c	ime	CMS e e c	ime
$\mathcal{S}_1$	100	52	84	11290
$\mathcal{S}_2$	100	85	5	1825
$\mathcal{S}_3$	100	145	72	4317
$\mathcal{S}_4$	9	4 0	5	554

Table 3. e c m a i f e D a d CEM^2 alg - i m a ame e e ima i acc ac .

Da a e	D	CEM^2
$\mathcal{S}_1$	0.0204	0.0204
$\mathcal{S}_2$	0.0171	0.0172
$\mathcal{S}_3$	0.03 3	0.03 3
$\mathcal{S}_4$	0.0308	0.0715

ig e a e f e D alg i m.

4.2. Unsupervised Classifications of Iris and Wine Data

e f e a l e D alg i m e e i ed cla i ca i f e I i a d i e da a f m UCI Mac i e ea i g e i 28 . e I i da a e c ai ee cla e , I i e ic l , I i gi ica a d I i Se a, a d eac cla c i f 50 am le . Eac am le i 4-dime i al ec mea i g e la m l g . I e e i- ime e I i da a, e e e i i ial al e f k a 6 a d e i i ial al e f e e a ame e a i im la- i e e ime . e e all , e D alg i m ed a $k^* = 3$ i e imal cla i ca i acc ac 9 .7 O l e f m 150 am le a e mi cla i ed . e e , i i ible a e D alg i m c e ge $k^* = 2$. Si ce e e a e I i b-cla e ic a e gl e- la ed, me li e a e al acce $k^* = 2$.

e i e da a e i 13-dime i al a d c i f 178 am le f ee i e . I i ca e, e e ce i da a e b e i ci al c m e a al i CA dime - i ed c i ec i e 29 a d c e l e ee i ci le c m e . e D alg i m i c d c ed e e ce ed da a i e i i ial al e f k a . E e ime al e l dem a e a e D alg i m al a c e ge i $k^* = 3$ a d e acc ac f cla i - ca i ca eac a 98.3 O l ee am le a e mi cla - i ed .

5. CONCLUSIONS

e a e i e iga ed e ela i i be ee e a- lea i g a d e M lea i g a d b idged em - i g a eg la i a i e m e a e age S a e f e e i babili e am le. a ed c a eg- la i a i mec a i m, e c c e d amicall eg- la i ed a m lea i g D f a ia mi e . c lli g e cale fac f i eg la i a i e m d amicall i c ea e f m 0 1, e D alg i m - a f m e a m lea i g i a ca abili f ada i e m del elec i , a d e g ad all a f m e c e i al ma im m li eli d lea i g b ai a c i e a ame e e ima i . M e e , e D al- g i m i calable a d ca be ed a big da a e i ce ai da a mma i a i ec i e. E e ime al e l dem a e a b e ic a d eal- ld da a e , e D alg i m ca l elec e c ec m- be f ac al a ia i a da a e b al b ai e M e ima e f e a ame e i e mi e.

Acknowledgments.

i a ed b e a al Scie ce da i f C i a f a 1171138 a d 07710 1.

6. REFERENCES

1. Da id eel a d Mac a la . i i e mi e m del , 2000.

2. ic a d A ed e a d me al e . Mi e de i ie , ma im m li eli d a d e em alg i m. *SIAM Review*, 2 2 195 239, 1984.

3. eil E Da . E ima i g e c m e f a mi e f mal di ib i . *Biometrika*, 5 3 4 3 474, 19 9.

4. A a iga . Di ib i blem i cl e i g. *Classification and Clustering*, age 45 72, 1977.

5. i g A ai e. A e l a e a i ical m del ide i ca i . *IEEE Transactions on Automatic Control*, 19 71 723, 1974.

ide Sc a e al. E ima i g e dime i f a m del. *The Annals of Statistics*, 2 4 1 4 4, 1978.

7. ma i a e . M deli g b e da a de c i - i . *Automatica*, 14 5 4 5 471, 1978.

8. C i S. allace a d Da id . D e. Mi im m me - age le g a d lm g c m le i . *The Computer Journal*, 42 4 270 283, 1999.

9. Mic ael D E c ba a d Mi e e a e ia de i- e ima i a d i fe e ce g mi e . *Journal of the American Statistical Association*, 90 430 577 588, 1995.

10. S l ia ic a d a d e e ee . O ba e ia a al i f mi e i a mbe f c m- e i di c i . *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59 4 731 792, 1997.

11. C a i C a i l a d A i idi i a . U e i ed lea i g f ga ia mi e ba ed a ia i al c m e li i g. *IEEE Transactions on Neural Networks*, 18 3 745 755, 2007.

12. Ma i A ig ei ed a d A il . ai . U e- i ed lea i g f i e mi e m del . *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 3 381 39 , 2002.

13. . M. C e g. Ma im m eig ed li eli d i- a i al e ali ed em f de i mi e cl e i g i a ma ic m del elec i . *IEEE Transactions on Knowledge and Data Engineering*, 17 705 7 1, 2005.

14. ei . i g- a g mac i e a ba e ia - llbac c eme f i ed lea i g a d e e l ec- a i a i . I *Proceedings of the 1995 International Conference on Neural Information Processing (ICONIP?5)*, l me 2, age 977 988, 1995.

15. ei . e a m , i ed cl a d a ma- ed m del elec i f e i ed a d e i ed lea i g ga ia mi e , ee-la e e a d me- bf- m m del . *International Journal of Neural Systems*, 11 01 43 9, 2001.

1 . ei . a m lea i g, c al cl, a d l gical elf- ga i i g mi e m del . *Neural Networks*, 15 8 1125 1151, 2002.

17. i e Ma, ai a g, a d ei . A g adie b a m lea i g e ga ia mi e i a - ma ed m del elec i . *Neurocomputing*, 5 481 487, 2004.

18. i e Ma a d e a g. a m lea i g i e mi e ada e g adie im leme a i a d a a i g cl mec a i m. *Neural Processing Letters*, 24 1 19 40, 200 .

19. i e Ma a d efe g e. A fa ed- i b a m lea i g alg i m ga ia mi e i a ma ed m del elec i . *Pattern Recognition Letters*, 29 701 711, 2008.

20. i e Ma. A ma ed m del elec i am - i e mi e a e e ical a al i . I *International Joint Conference on Neural Networks 2006 (IJCNN'06)*, age 4139 4145. IEEE, 200 .

21. ei . a e ia i g a g mac i e, cl e i g a d mbe f cl e . *Pattern Recognition Letters*, 18 11 11 7 1178, 1997.

22. i a d i e Ma. A g adie e eg la - i ed li eli d lea i g alg i m ga ia mi e i a ma ic m del elec i . I *Advances in Neural Networks-ISNN 2006*, age 4 4 4 9. S i ge , 200 .

23. i a d ace -S i g I . e e ali ed c m- e i i e lea i g f ga ia mi e m del . *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 39 4 901 909, 2009.

24. ei , Adam a , a d E i O a. i al e- ali ed c m e i i e lea i g f cl e i g a al i , bf e a d c e de ec i . *Neural Networks, IEEE Transactions on*, 4 4 3 49, 1993.

25. i e Ma a d ai a g. A c -f c i a - ac i al e ali ed c m e i i e lea i g cl . *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 3 4 722 737, 200 .

2 . id g i , Ma - e g g, a d . S. e g. Scalable m del-ba ed cl e i g f la ge da aba - e ba ed da a mma i a i . *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 11 1710 1719, 2005.

27. i e Ma a d ia fe g i . e b a eali g lea - i g alg i m f ga ia mi e i a ma ed m del elec i . *Pattern Recognition*, 40 7 2029 2037, 2007.

28. UCI Mac i e ea i g e i . //.ic . ci.ed /∼mlea /.

29. Ia lliffe. *Principal Component Analysis*. ile O li e ib a , 2005.