한국통신학회논문지 '13-02 Vol.38C No.02
http://dx.doi.org/10.7840/kics.2013.38C.2.208

라그랑지 기법을 쓴 영 공간 기반 선형 판별 분석법의 변형 기법

호우위시◆, 민 황 기*, 송 익 호°, 최 명 수**, 박 선**, 이 성 로***

Transformation Technique for Null Space-Based Linear Discriminant Analysis with Lagrange Method

Yuxi Hou◆, Hwang-Ki Min*, Iickho Song°, Myeong Soo Choi**, Sun Park**, Seong Ro Lee***

요 약

부류안 분산 행렬의 특이성 때문에 선형 판별 분석은 작은 표본 크기 문제에 쓰기에 알맞지 않다. 이에 선형 판별 분석을 확장하여 작은 표본 크기 문제에서 좋은 성능을 갖는 영 공간 기반 선형 판별 분석이 제안되었다. 이 논문에서는 라그랑지 기법을 바탕으로 하여, 영 공간 기반 선형 판별 분석을 써서 특징을 추출하는 문제를 선형 방정식 문제로 바꾸는 과정을 제안하였다.

Key Words : feature extraction, Lagrange method, null space-based linear discriminant analysis

ABSTRACT

Due to the singularity of the within-class scatter, linear discriminant analysis (LDA) becomes ill-posed for small sample size (SSS) problems. An extension of LDA, the null space-based LDA (NLDA) provides good discriminant performances for SSS problems. In this paper, by applying the Lagrange technique, the procedure of transforming the problem of finding the feature extractor of NLDA into a linear equation problem is derived.

Ⅰ. 서 론

선형 판별 분석법은 (linear discriminant analysis: 줄여서, 선판분)[1] 패턴 인식 문제에서 자료의 특징을 추출하는 방법들 가운데 가장 널리 쓰이는 방법의 하나이다. 하지만 자료의 차원 d가 자료의 수 N보다 큰 패턴 인식 문제에서는 부류안 분산 행렬의 (within-class scatter) 특이성 (singularity) 때문에 선판분을 쓸 수 없다는 문제가 있다. 여기서, 부류의 개수를 c라 할 때 $d>N-c$인 패턴 인식 문제를 작은 표본 크기 (small sample size) 문제라고 부른다.

작은 표본 크기 문제에서도 선판분을 쓸 수 있도록 하기 위해서, 선판분을 바탕으로 하여 몇 가지

※ 이 논문은 교육과학기술부의 재원으로 한국연구재단이 선정하여 지원하는 중견연구자지원사업 2012-0005622와, 지식경제부의 재원으로 정보통신산업진흥원이 지원하는 대학 IT연구센터 지원사업 과제 NIPA-2012-H0301-12-2005를 수행하여 얻은 결과 가운데 하나입니다.
◆ 주저자 : 한국과학기술원 전기 및 전자공학과 통계학적 신호처리 연구실, hyyyx@Sejong.kaist.ac.kr, 정회원
° 교신저자 : 한국과학기술원 전기 및 전자공학과 통계학적 신호처리 연구실, isong@Sejong.kaist.ac.kr, 종신회원
* 한국과학기술원 전기 및 전자공학과 통계학적 신호처리 연구실, hkmin@Sejong.kaist.ac.kr
** 목포대학교 정보산업연구소, mschoi@mokpo.ac.kr, 정회원, sunpark@mokpo.ac.kr, 정회원
*** 목포대학교 정보전자공학과, srlee@mokpo.ac.kr, 정회원
논문번호 : KICS2013-01-035, 접수일자 : 2013년 1월 14일, 최종논문접수일자 : 2013년 2월 19일

관련 기법들이 제안된 바 있다: 주성분 분석에 (principle component analysis) 선판분을 응용하여 더한 방법[2], 직접 선판분 (direct linear discriminant analysis)[3], 영 공간 기반 선판분 (null space-based linear discriminant analysis: 줄여서, 영선판분)[4,5], 직교 선판분 (orthogonal linear discriminant analysis)[6], 그리고 주파수 회귀 판별 분석법이 (spectral regression discriminant analysis)[7] 대표적인 방법들이다. 이 가운데, 영선판분 방법은 일반적으로 좋은 패턴 인식 성능을 가지고 있지만 그 복잡도가 높다는 문제가 있다. 영선판분을 써서 특징을 추출하는 문제를 더 간단히 풀고자, 큐알 인수분해 (QR factorization) 기법과[8] 고유값 분해 (eigen-decomposition) 기술을 바탕으로 한 기법이[9] 각각 제안된 바 있다.

이 논문에서는 라그랑지 (Lagrange) 기법을 써서, 영선판분을 써서 특징을 추출하는 문제를 선형 방정식 문제로 바꾸는 과정을 보인다.

Ⅱ. 영 공간 기반 선형 판별 분석법

패턴 인식에서, c 부류 패턴 인식 문제는 일대다 (one-versus-all) 방법으로[10], 두 부류 패턴 인식 문제 c개로 바꿀 수 있기 때문에, 이 논문에서는 두 부류 패턴 인식 문제만을 다룰 것이다. 선형 특징 추출기는 $F(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{x} + b$로 널리 표현된다. 여기서, $\boldsymbol{x} \in R^d$는 입력이고, $d \times 1$ 벡터 $\boldsymbol{w}$는 선형 판별 벡터라고 (linear discriminant vector: 줄여서, 선판벡) 부르며, 스칼라 b는 바이어스이다.

이제, 자료 $\boldsymbol{x}_i \in R^d$와 그 표식 $y_i \in \{1, -1\}$들의 집합을 $D = \{\boldsymbol{x}_i, y_i\}_{i=1}^N$이라 가정하자. 문제를 간단히 하기 위해, 자료 $\{\boldsymbol{x}_i\}$가 정렬되어 있어서 표식 벡터 $\boldsymbol{y} = [y_1\ y_2 \cdots\ y_N] = \left[\mathbf{1}_{1\times N_1} - \mathbf{1}_{1\times N_2}\right]^T$처럼 쓸 수 있다고 하자. 이때, $\mathbf{1}_{a\times b}$는 모든 원소가 1인 $a \times b$ 행렬이다.

영선판분은, 제한조건이 $\boldsymbol{w}^T\boldsymbol{S}_W\boldsymbol{w} = 0$ 일 때 선판벡

$$\boldsymbol{w}_N = \arg\max_{\|\boldsymbol{w}\|=1} \boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w} \tag{1}$$

을 얻는 문제이다. 여기서, ‖ • ‖는 유클리드 길이이고 (Euclidean norm),

$$\boldsymbol{S}_B = (\boldsymbol{m}_1 - \boldsymbol{m}_2)(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T \tag{2}$$

는 차수가 (rank) $c-1$인 $d \times d$ 부류사이 분산행렬이며 (between-class scatter),

$$\boldsymbol{S}_W = \boldsymbol{X}\boldsymbol{X}^T - N_1\boldsymbol{m}_1\boldsymbol{m}_1^T - N_2\boldsymbol{m}_2\boldsymbol{m}_2^T \tag{3}$$

은 차수가 $\min(d, N-c)$인 $d \times d$ 부류안 분산 행렬이다. 여기서,

$$\boldsymbol{X} = [\boldsymbol{x}_1\ \boldsymbol{x}_2 \cdots\ \boldsymbol{x}_N] \tag{4}$$

는 $d \times N$ 자료 행렬이고, $Z_l = \{i : x_i \in \text{부류 } l\}$은 부류 l의 가리킴 수 집합이며,

$$\boldsymbol{m}_l = \frac{1}{N_l}\sum_{\{i \in Z_l\}} \boldsymbol{x}_i \tag{5}$$

는 부류 l의 표본 평균 벡터이다. 또한, N_l은 Z_l의 원소들의 개수이다. 일단 선판벡 $\boldsymbol{w}_N$이 정해지면, 바이어스 b_N은

$$b_N = \frac{1}{N}\sum_{i=1}^{N}(y_i - \boldsymbol{w}_N^T\boldsymbol{x}_i) \tag{6}$$

으로 얻을 수 있다.

Ⅲ. 라그랑지 기법을 써서 선형 방정식 문제로 바꾸기

먼저, $\boldsymbol{w}^T\boldsymbol{w} = \|\boldsymbol{w}\|^2 = 1$ 에서

$$\frac{1}{\boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w}} = \frac{\boldsymbol{w}^T\boldsymbol{w}}{\left\{(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T\boldsymbol{w}\right\}^T(\boldsymbol{m}_1 - \boldsymbol{m}_2)^T\boldsymbol{w}} \tag{7}$$

이기 때문에, (1)을

$$\gamma_H = \arg\min_{\gamma} \frac{1}{4}\gamma^T\gamma \tag{8}$$

로 다시 쓸 수 있다. 여기서,

$$\gamma = \frac{2w}{(m_1 - m_2)^T w} \tag{9}$$

이고

$$\gamma_H = \frac{2w_N}{(m_1 - m_2)^T w_N} \tag{10}$$

이다. 또한, 제한조건 $w^T S_W w = 0$을

$$X^T \gamma + b1_{N\times 1} = y \tag{11}$$

로 다시 쓸 수 있다. 여기서,

$$b = -\frac{(m_1 + m_2)^T w}{(m_1 - m_2)^T w} \tag{12}$$

이다. 이제, (8)과 (11)에 나타난 문제의 라그랑지 함수를

$$L_1 = \frac{1}{4}\gamma^T \gamma + \alpha^T (X^T \gamma + b1_{N\times 1} - y) \tag{13}$$

으로 둘 수 있다. 이때, $N\times 1$ 벡터 α는 라그랑지 곱수이다 (multiplier). 라그랑지 함수 L_1을 b와 γ에 대해서 각각 편미분한 값이 0이 되도록 만들면,

$$1_{1\times N}\alpha = 0 \tag{14}$$

와

$$\gamma = -2X\alpha \tag{15}$$

를 얻을 수 있다.

이제 (14)와 (15)를 써서, 라그랑지 함수 (13)을

$$L_1 = -\alpha^T (X^T X\alpha + y) \tag{16}$$

으로 다시 쓸 수 있다. 따라서, 제한조건이 (11)인 문제 (8)을 제한조건이 (14)인 문제

$$\alpha_H = \arg\min_{\alpha} \alpha^T (X^T X\alpha + y) \tag{17}$$

로 다시 쓸 수 있고, 이때 α_H와 γ_H의 관계는

$$\gamma_H = -2X\alpha_H \tag{18}$$

임을 (15)에서 알 수 있다. 라그랑지 기법을 써서 앞과 비슷하게 몇 단계 거치고 나면, 제한조건이 (14)인 문제 (17)을

$$\begin{aligned}\alpha_H = \arg\min_{\alpha} \Big(&\alpha^T X^T X\alpha + \alpha^T y - \\ &\frac{2}{N} 1_{1\times N}\alpha 1_{1\times N} X^T X\alpha - \\ &\frac{1}{N} 1_{1\times N}\alpha 1_{1\times N} y\Big)\end{aligned} \tag{19}$$

처럼 쓸 수 있다.

식 (19)에 보인 목적함수를 α에 대해 미분한 값을 0으로 두면,

$$A_L \alpha_H = y_H \tag{20}$$

임을 알 수 있다. 여기서,

$$A_L = X^T X - \frac{1}{N} 1_{N\times N} X^T X - \frac{1}{N} X^T X 1_{N\times N} \tag{21}$$

이고,

$$y_H = \frac{1}{2}\left(\frac{1}{N} 1_{N\times N} - I_N\right) y = -\frac{N_1 N_2}{N^2}\bar{y} \tag{22}$$

이다. 여기서, I_N은 $N\times N$ 단위행렬이고,

$$\bar{y} = \left[\frac{N}{N_1} 1_{1\times N_1} \quad -\frac{N}{N_2} 1_{1\times N_2}\right]^T \tag{23}$$

은 응답벡터이다. 이제, (10)과 (18)에서

$$w_N = -(m_1 - m_2)^T w_N X\alpha_H \tag{24}$$

이고, (1)에서 $\| w_N \| = 1$이기 때문에 스칼라 $(m_1 - m_2)^T w_N$의 값이 $\pm\frac{1}{\| X\alpha_H \|}$임을 알 수 있다. 그러므로,

$$w_N = \pm \frac{X\alpha_H}{\| X\alpha_H \|} \tag{25}$$

이다. 곧, 선형 방정식 (20)를 풀어 α_H를 얻으면

선판벡 w_N을 얻을 수 있다.

벡터 $m_1 - m_2$가 S_W의 영 공간과 직교하지 않으며, $X^T X$가 완전 계수 행렬이라는 조건에서 (20)를 얻었음을 새겨두자. 한편, $m_1 - m_2$가 S_W의 영 공간과 직교할 때는,

$$w^T S_B w = \left|(m_1 - m_2)^T w\right|^2 \tag{26}$$

이 늘 0이기 때문에 (1)의 풀이가 없다. 다만, $X^T X$가 특이행렬일 때에는, 작은 양수 μ를 써서 $X^T X$를 완전 계수 행렬 $X^T X + \mu I_N$로 바꾸는 방법을 (섭동법) 써서 α_H를 얻을 수 있다.

Ⅳ. 벡터 α_H 얻기

식 (20)을 써서 바로 α_H를 얻는 방법은 $\frac{4}{3}N^3 + O(N^2)$쯤의 곱셈이 든다[11]. 계산량을 줄이기 위해, 바로 얻는 방법 대신에 다음 방법으로 α_H를 얻어보자. 식 (14)에서 $1_{1\times N}\alpha_H = 0$이므로, 행렬

$$\widetilde{A}_L = A_L + \mu_H 1_{N\times N} \tag{27}$$

이

$$\widetilde{A}_L \alpha_H = y_H \tag{28}$$

을 만족시킴을 쉽게 알 수 있다. 이제, $m^T m$보다 큰 실수 μ_H를 생각하면 행렬 $\widetilde{A}_L$은 양정치이기 때문에, 촐레스키 분해를 써서 (28)을 풀면 α_H의 해를 얻을 수 있다. 이렇게 α_H를 얻으면 직접 얻는 것 보다 복잡도가 더 낮을 것이라는 것을 예상할 수 있다.

Ⅴ. 맺음말

이 논문에서는 라그랑지 기법을 써서, 영 공간 기반 선형 판별 분석법을 써서 특징을 추출하는 문제를 선형 방정식 문제로 바꾸었다. 이렇게 얻어진 선형 방정식 문제를 풀어 특징 추출기를 얻는 과정은, 원래의 문제를 풀어 특징 추출기를 얻는 과정보다 복잡도가 더 낮을 것으로 기대한다.

참고문헌

[1] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed., John Wiley and Sons, 2001.

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Patt. Anal., Mach. Intell.*, vol. 19, no. 7, pp. 711-720, July 1997.

[3] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data – with application to face recognition," *Patt. Recogn.*, vol. 34, no. 10, pp. 2067-2070, Oct. 2001.

[4] L.-F. Chen, H.-Y. Mark Liao, M.-T. Ko, J.-C. Lin, and G.-J. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Patt. Recogn.*, vol. 33, no. 10, pp. 1713-1726, Oct. 2000.

[5] D.-U. Cho, U.-D. Chang, Y.-G. Kim, Y.-J. Song, J.-H. Ahn, and B.-H. Kim, "2D direct LDA algorithm for face recognition," *J. Korean Inform., Comm. Soc.*, vol. 30, no. 12C, pp. 1162-1166, Dec. 2005.

[6] J. Ye and T. Xiong, "Computational and theoretical analysis of null space and orthogonal linear discriminant analysis," *J. Mach. Learn. Res.*, vol. 7, no. 7, pp. 1183-1204, July 2006.

[7] D. Cai, X. He, and J. Han, "SRDA: An efficient algorithm for largescale discriminant analysis," *IEEE Trans. Knowl., Data Eng.*, vol. 20, no. 1, pp. 1-12, Jan. 2008.

[8] D. Chu and G. S. Thye, "A new and fast implementation for null space based linear discriminant analysis," *Patt. Recogn.*, vol. 43, no. 4, pp. 1373-1379, Apr. 2010.

[9] R. Huang, Q. Liu, H. Lu, and S. Ma, "Solving the small sample size problem of LDA," *Proc. Int. Conf. Patt. Recogn.*, pp. 29-32, Quebec, Canada, Aug. 2002.

[10] R. Rifkin and A. Klautau, "In defense of one-vs-all classification," *J. Mach. Learn.*

Res., vol. 5, no. 1, pp. 101-141, Jan. 2004.
[11] G. W. Stewart, *Matrix Algorithms Volume I: Basic Decompositions*, Society for Industrial and Applied Mathematics, 1998.

호우위시 (Yuxi Hou)

2001년 2월 시안교통대학교 전기 및 전자공학과 공학사
2005년 2월 시안교통대학교 전기 및 전자공학과 공학석사
20012년 8월 한국과학기술원 전기 및 전자공학과 공학박사
2012년 8월~현재 한국과학기술원 전기 및 전자공학과 박사후 연구과정
<관심분야> 특징 추출, 패턴 인식, 신호처리, 기계학습

민 황 기 (Hwang-Ki Min)

2004년 2월 한국과학기술원 전기 및 전자공학과 공학사
2006년 2월 한국과학기술원 전기 및 전자공학과 공학석사
2006년 3월~현재 한국과학기술원 전기 및 전자공학과 박사과정
<관심분야> 신경회로망, 기계 학습, 패턴 인식

송 익 호 (Iickho Song)

1982년 2월, 1984년 2월 서울대학교 전자공학과 공학사(준최우등), 공학석사
1985년 8월, 1987년 5월 펜실베니아대학교 전기공학과 공학석사, 공학박사
1987년 3월~1988년 2월 벨 통신연구소 연구원
1988년 3월~현재 한국과학기술원 전기 및 전자공학과 조교수, 부교수, 교수
1995년 1월~현재 한국통신학회 논문지 편집위원
대한전자공학회, 한국음향학회, 한국통신학회 평생회원, IET 석학회원, IEEE 석학회원
<관심분야> 통계학적 신호처리와 통신이론, 신호검파와 추정, 이동통신

최 명 수 (Myeong Soo Choi)

2000년 2월 목포대학교 전자공학과 공학사
2002년 2월 목포대학교 전자공학과 공학석사
2009년 2월 목포대학교 전자공학과 공학박사
2009년 3월 목포대학교 해양텔레매틱스기술개발센터 박사후연구원
2009년 12월~현재 목포대학교 정보산업연구소 연구전임교수
<관심분야> 디지털통신시스템, USN, 배열신호처리, 임베디드시스템

박 선 (Sun Park)

1996년 2월 전주대학교 전자계산학과 이학사
2001년 2월 한남대학교 정보통신학과 공학석사
2007년 2월 인하대학교 컴퓨터정보공학과 공학박사
2008~2009년 호남대학교 컴퓨터공학과 전임강사
2010년 전북대학교 인력양성사업단 박사후 과정
2010년 12월~현재 목포대학교 정보산업연구소 연구전임교수
<관심분야> 정보검색, 데이터마이닝, 데이터베이스, 해양IT정보융합디지털통신시스템, 이동 및 위성통신시스템, USN/텔레미틱스응용분야, 임베디드시스템

이 성 로 (Seong Ro Lee)

1987년 2월 고려대학교 전자공학과 공학사
1990년 2월 한국과학기술원 전기및전자공학과 공학석사
1996년 8월 한국과학기술원 전기및전자공학과 공학박사
1997년 9월~현재 목포대학교 공과대학 정보전자공학과 교수
<관심분야> 디지털통신시스템, 이동 및 위성통신시스템, USN/텔레매틱스응용분야, 임베디드시스템