

Discussion: Robust Sparse Quadratic Discrimination

Han Xiao

Department of Statistics & Biostatistics
Rutgers University

2014 Rutgers Statistics Symposium
Statistics and the Century of Data

May 02 2014

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.
- **QUADRO** provides a nice solution featuring:

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.
- **QUADRO** provides a nice solution featuring:
 - Criterion: Rayleigh quotient

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.
- **QUADRO** provides a nice solution featuring:
 - Criterion: Rayleigh quotient
 - Avoid the fourth order joint moments, under elliptical distributions.

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.
- QUADRO provides a nice solution featuring:
 - Criterion: Rayleigh quotient
 - Avoid the fourth order joint moments, under elliptical distributions.
 - Non-Gaussian, non-equal covariance matrices.

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.
- QUADRO provides a nice solution featuring:
 - Criterion: Rayleigh quotient
 - Avoid the fourth order joint moments, under elliptical distributions.
 - Non-Gaussian, non-equal covariance matrices.
 - Convex optimization.

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.
- **QUADRO** provides a nice solution featuring:
 - Criterion: Rayleigh quotient
 - Avoid the fourth order joint moments, under elliptical distributions.
 - Non-Gaussian, non-equal covariance matrices.
 - Convex optimization.
 - Robust estimation of mean and covariance matrix.

High Dimensional Classification

- Quadratic discrimination has been a very challenging problem under high dimensionality.
- **QUADRO** provides a nice solution featuring:
 - Criterion: Rayleigh quotient
 - Avoid the fourth order joint moments, under elliptical distributions.
 - Non-Gaussian, non-equal covariance matrices.
 - Convex optimization.
 - Robust estimation of mean and covariance matrix.
 - Theoretical guarantee of the performance.

Gaussian Quadratic Discriminant Analysis

QDA (Gaussian) compares discriminant functions

$$f_k(\mathbf{X}) = -\frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (\mathbf{X} - \boldsymbol{\mu}_k)^\top \Sigma_k^{-1} (\mathbf{X} - \boldsymbol{\mu}_k) + \log \pi_k.$$

- Robust estimation of means as in QUADRO.
- Sparse and robust estimation of Σ_k^{-1} . (*Meinshausen & Buhlmann, 06, Yuan & Lin, 07, Lam & Fan, 09, Cai et al, 11, Liu et al, 12, Xue & Zou, 12.*)

The elliptical distribution has the density

$$|\Sigma_k|^{-1/2} g_k [(\mathbf{x} - \boldsymbol{\mu}_k)^\top \Sigma_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k)].$$

- Estimate g_k and apply the Bayes rule?

Regularized Inputs to QUADRO

QUADRO uses robust estimates of Σ_k and μ_k as inputs.

- In LDA, can bypass the covariance matrix estimation by estimating the coefficients of the linear discriminant function directly. (*Cai et al, 11, Fan et al, 12, Mai et al, 12., Han et al, 13.*)
- Is there any benefit of exploiting the structure of Σ_k ?

Two Stage Procedure

When d is very large, can take an additional screening step.

- Sure screen under regression setting. (*Fan & Lv, 08, Fan & Song, 10, Fan et al, 11.*)
- Features annealed independence rule. (*Fan & Fan, 08.*)
- Hierarchical structure? (*McCullagh & Nelder, 89, Hamada & Wu, 92, Yuan et al, 09, Zhao et al, 09, Choi et al, 10, Bien et al, 13.*)

Thank You!