Some results on Imprecise discriminant analysis

11th Workshop on Principles and Methods of Statistical Inference with Interval Probability

CARRANZA-ALARCON Yonatan-Carlos
Ph.D. Candidate in Computer Science

DESTERCKE Sébastien
Ph.D Director

30 July 2018 to 01 August 2018

Overview

Imprecise Discriminant Analysis Classification

- Classification
  - Decision Making
  - Discriminant Analysis
- Imprecise Classification
  - Imprecise Decision
  - Imprecise Linear discriminant analysis
- Future work
- Conclusions

Overview

- Classification
  - Decision Making
  - Discriminant Analysis
- Imprecise Classification
  - Imprecise Decision
  - Imprecise Linear discriminant analysis
- Future work
- Conclusions

heudiasyc

Classification - Setting

A classic classification problem is composed of :

- Data training $D = \{x_i, y_i\}_{i=0}^{N}$ such as :
  - (Input) $x_i \in \mathscr{X}$ are regressors or features (often $x_i \in \mathbb{R}^p$).
  - (Output) $y_i \in \mathscr{K}$ is a response category variable, with $\mathscr{K} = \{ m_1, ..., m_K\}$

Classification - Setting

A classic classification problem is composed of :

- Data training $D = \{x_i, y_i\}_{i=0}^{N}$ such as :
  - (Input) $x_i \in \mathscr{X}$ are regressors or features (often $x_i \in \mathbb{R}^p$).
  - (Output) $y_i \in \mathscr{K}$ is a response category variable, with $\mathscr{K} = \{ m_1, ..., m_K\}$

Objective

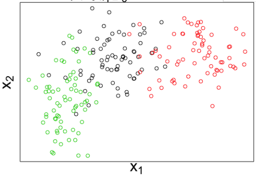
Given training data $D = \{x_i, y_i\}_{i=0}^{N}$, we need to learn a classification rule : $\phi : \mathscr{X} \rightarrow \mathscr{Y}$ in order to predict a new observation $\phi(\boldsymbol{x}^*)$

heudiasyc

Classification - Outline (Example)

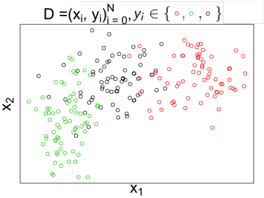
Getting Training Data

$D = (x_i, y_i)_{i=0}^{N}, y_i \in \{ \circ, \circ, \circ \}$



Classification - Outline (Example)

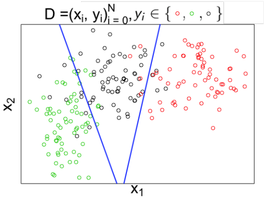
Getting Training Data



→

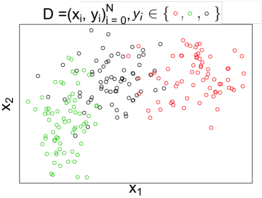
Learning a classification rule :

$$\phi : \mathscr{X} \rightarrow \mathscr{Y}$$



Classification - Outline (Example)

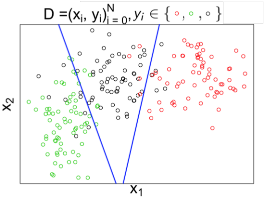
Getting Training Data



→

Learning a classification rule :

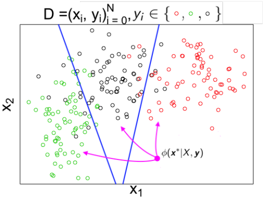
$$\phi : \mathscr{X} \rightarrow \mathscr{Y}$$



→

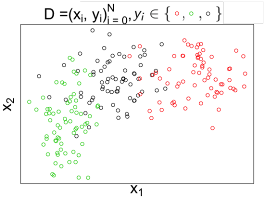
Predict class for new instances :

$$\widehat{y}^* := \phi(\boldsymbol{x}^* | X, \boldsymbol{y})$$



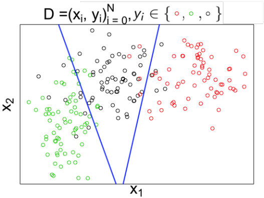
Classification - Outline (Example)

Getting Training Data



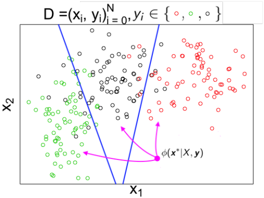
Learning a classification rule :

$$\phi : \mathscr{X} \rightarrow \mathscr{Y}$$



Predict class for new instances :

$$\widehat{y}^* := \phi(\boldsymbol{x}^* | X, \boldsymbol{y})$$



But :

- How can we learn the *“classification rule”* (model) from training data ?

Decision Making in Statistic

- In statistic : classification rule often seen as a decision-making problem under risk of getting missclassification.

$$\mathscr{R}(y,\varphi(X)) = \underset{\varphi(X)\in\mathscr{K}}{\arg\min}\ \mathbb{E}_{\mathscr{X}\times\mathscr{Y}}[\mathscr{L}(y,\varphi(X))] \quad (1)$$

- Under $1/0$ loss function $\mathscr{L}$, minimizing $\mathscr{R}$ equivalent to :

$$\phi(\boldsymbol{x}^*|X,\boldsymbol{y}) := \underset{m_k\in\mathscr{K}}{\arg\max}\ P(y = m_k|X = \boldsymbol{x}^*) \quad (2)$$

- Where :
  1. The predicted class $\hat{y}^* = \phi(\boldsymbol{x}^*|X,\boldsymbol{y})$ is the most probable (equation (2)).
  2. This last equation (2) is also known as Bayes classifier [1, pp. 21].

Decision Making in Statistic

- In statistic : classification rule often seen as a decision-making problem under risk of getting missclassification.

$$\mathscr{R}(y,\varphi(X)) = \underset{\varphi(X)\in\mathscr{K}}{\arg\min}\ \mathbb{E}_{\mathscr{X}\times\mathscr{Y}}\left[\mathscr{L}(y,\varphi(X))\right] \quad (1)$$

- Under $1/0$ loss function $\mathscr{L}$, minimizing $\mathscr{R}$ equivalent to :

$$\phi(\boldsymbol{x}^*|X,\boldsymbol{y}) := \underset{m_k\in\mathscr{K}}{\arg\max}\ P(y = m_k|X = \boldsymbol{x}^*) \quad (2)$$

- Where :
  1. The predicted class $\hat{y}^* = \phi(\boldsymbol{x}^*|X,\boldsymbol{y})$ is the most probable (equation (2)).
  2. This last equation (2) is also known as Bayes classifier [1, pp. 21].

Decision Making in Statistic

Definition (Preference ordering [5, pp. 47])

With general loss $\mathscr{L}(\cdot,\cdot)$, m_a is preferred to m_b, denoted by $m_a \succ m_b$, if and only if :

$$\mathbb{E}_P[\mathscr{L}(\cdot, m_a)|\boldsymbol{x}^*] < \mathbb{E}_P[\mathscr{L}(\cdot, m_b)|\boldsymbol{x}^*]$$

In the particular case where $\mathscr{L}(\cdot,\cdot)$ is the $0/1$ loss function we get :

$$m_a \succeq m_b \iff \frac{P(y = m_a|X = \boldsymbol{x}^*)}{P(y = m_b|X = \boldsymbol{x}^*)} > 1$$

where $P(Y = m_a|X = \boldsymbol{x}^*)$ is the class probability. We then take the ***maximal element*** of the complete order $\succeq$, i.e.

$$m_{i_K} \succeq m_{i_{K-1}} \succeq .... \succeq m_{i_1} \iff P(y = m_{i_K}|\boldsymbol{x}^*) \geq .... \geq P(y = m_{i_1}|\boldsymbol{x}^*)$$

(Precise) Discriminant Analysis

Applying Baye's rules to $P(Y = m_a | X = \boldsymbol{x}^*)$:

$$P(y = m_k | X = \boldsymbol{x}^*) = \frac{P(X = \boldsymbol{x}^* | y = m_k) P(y = m_k)}{\sum_{m_l \in \mathcal{K}} P(X = \boldsymbol{x}^* | y = m_l) P(y = m_l)}$$

where $\pi_k := \mathbb{P}_{Y=y_k}$ such as $\sum_j^K \pi_j = 1$ and $\mathcal{G}_k := \mathbb{P}_{X|Y=m_k} \sim \mathcal{N}(\mu_k, \Sigma_k)$

A frequentist point estimation :

$$\widehat{\pi}_k = \frac{n_k}{N}$$

$$\widehat{\mu}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} x_{i,k}$$

$$\widehat{\Sigma}_k = \frac{1}{N - n_k} \sum_{i=1}^{n_k} (x_{i,k} - \overline{\boldsymbol{x}}_k)(x_{i,k} - \overline{\boldsymbol{x}}_k)^t$$

Overview

- Classification
  - Decision Making
  - Discriminant Analysis
- Imprecise Classification
  - Imprecise Decision
  - Imprecise Linear discriminant analysis
- Future work
- Conclusions

Decision Making in Imprecise Probabilities

Definition (Partial Ordering by Maximality Criterion)

Let $\mathscr{P}$ a set of probabilities, then m_a is preferred to m_b if the cost of exchanging m_a with m_b have a positive lower expectation :

$$m_a \succ_M m_b \iff \inf_{P \in \mathscr{P}} \mathbb{E}_P[\mathscr{L}(\cdot, m_b) - \mathscr{L}(\cdot, m_a) | \boldsymbol{x}^*] > 0$$

if $\mathscr{L}(\cdot,\cdot)$ is $1/0$ loss function, so :

$$m_a \succ_M m_b \iff \inf_{P \in \mathscr{P}} \frac{P(y = m_a | X = \boldsymbol{x}^*)}{P(y = m_b | X = \boldsymbol{x}^*)} > 1$$

Decision Making in Imprecise Probabilities

By applying Bayes theorem on $P(y = m_a | X = \boldsymbol{x}^*)$, so :

$$m_a \succ_M m_b \iff \inf_{P_{X|y} \in \mathscr{P}_1, P_y \in \mathscr{P}_2} \frac{P(\boldsymbol{x}^* | y = m_a) P(y = m_a)}{P(\boldsymbol{x}^* | y = m_b) P(y = m_b)} > 1$$

The resulting set of cautions decisions is :

$$Y_M = \{m_a \in \mathscr{K} \mid \nexists m_b : m_a \succ_M m_b\}$$

For instance, if $\mathscr{K} = \{m_a, m_b, m_c\}$, we can have :

$$\widehat{Y}_M = \{m_a \succ_M m_b, m_c \succ_M m_b, m_a \succ\prec_M m_c\} = \{m_a, m_c\}$$

Imprecise Linear Discriminant Analysis (ILDA)

Objective :

Making imprecise the parameter mean μ_k of each Gaussian distribution family $\mathscr{G}_k := \mathbb{P}_{X|Y=m_k} \sim \mathscr{N}(\mu_k, \widehat{\Sigma})$

Assumptions :

- Covariances precisely estimated and Homoscedasticity, i.e. $\Sigma_k = \Sigma$:

$$\widehat{\Sigma} = \frac{1}{(N-K)} \sum_{k=1}^{K} \sum_{i=1}^{n_k} (x_{i,k} - \overline{\boldsymbol{x}}_k)(x_{i,k} - \overline{\boldsymbol{x}}_k)^t$$

- Prior probabilities precisely estimated : $\widehat{\pi}_k = \frac{n_k}{N}$

Decision Making in ILDA

We take the previously maximality criterion and assumptions, so :

$$m_a \succ_M m_b \iff \inf_{P_{X|y} \in \mathscr{P}_1, P_y \in \mathscr{P}_2} \frac{P(\boldsymbol{x}^* | y = m_a) P(y = m_a)}{P(\boldsymbol{x}^* | y = m_b) P(y = m_b)} > 1 \qquad (3)$$

$$\iff \inf_{P_{X|y} \in \mathscr{P}_1} \frac{P(\boldsymbol{x}^* | y = y_a) \ \ \widehat{\pi}_a}{P(\boldsymbol{x}^* | y = y_b) \ \ \widehat{\pi}_b} > 1$$

Given $\mathscr{G}_k := \mathbb{P}_{X|Y=m_k} \sim \mathscr{N}(\mu_k, \widehat{\Sigma})$ are independent :

$$\iff \frac{\inf_{P \in \mathscr{G}_a} P(\boldsymbol{x}^* | y = y_a) \ \ \widehat{\pi}_a}{\sup_{P \in \mathscr{G}_b} P(\boldsymbol{x}^* | y = y_b) \ \ \widehat{\pi}_b} > 1$$

Decision Making in ILDA (cont...)

Then, the problem reduces to two optimisation problems :

$$\underline{P}(\boldsymbol{x}^*|y=y_a) = \inf_{P\in\mathscr{G}_a} P(\boldsymbol{x}^*|y=y_a) \quad (4)$$

$$\overline{P}(\boldsymbol{x}^*|y=y_b) = \sup_{P\in\mathscr{G}_b} P(\boldsymbol{x}^*|y=y_b) \quad (5)$$

As $\mathbb{P}_{X|Y=m_k} \sim \mathcal{N}(\mu_k, \widehat{\Sigma})$ and $\Sigma_b = \widehat{\Sigma}$, so :

$$\underline{P}(\boldsymbol{x}^*|y=y_a) \iff \underline{\mu}_a = \inf_{P\in\mathscr{G}_a} -\frac{1}{2}(\boldsymbol{x}^* - \mu_a)^T\widehat{\Sigma}^{-1}(\boldsymbol{x}^* - \mu_a) \quad (6)$$

$$\overline{P}(\boldsymbol{x}^*|y=y_b) \iff \overline{\mu}_b = \sup_{P\in\mathscr{G}_b} -\frac{1}{2}(\boldsymbol{x}^* - \mu_b)^T\widehat{\Sigma}^{-1}(\boldsymbol{x}^* - \mu_b) \quad (7)$$

Imprecise Linear Discriminant Analysis

Now, the question is : How could we make imprecise the unknown mean parameter μ_k ?

- Confidence intervals.
- Neighbors around μ_k.
- P-Box
- Robust Bayesian
- ....

*We would use **robust Bayesian** with **conjugate distributions for exponential families***

Imprecise Linear Discriminant Analysis
Bayesian inference context

In classic Bayesian inference is based on two components :

- The distribution of the observed data conditional on its unknown parameters (or Likelihood).
- A belief information of expert (or prior distribution).

In order to build procedures of posterior inference on the unknown parameter, in this case μ_k.

$$p(\mu_k \mid X, \boldsymbol{y} = m_k) \propto p(X \mid \mu_k, \boldsymbol{y} = m_k) p(\mu_k) \tag{8}$$

Where $p(\mu_k) \in \mathscr{P}_{\mu_k}$ could belong a set of prior distributions $\mathscr{P}_{\mu_k}$

Imprecise Linear Discriminant Analysis

We propose to use a set of prior distributions based on near-ignorance approach of [6, eq. 16] :

$$\mathcal{M}_0^{\mu} = \left\{ \mu \in \mathbb{R}^d \middle| p(\mu|m) \propto \exp(\ell^T \mu),\ m = [\ell_1, ..., \ell_d]^T \in \mathbb{L} \right\} \quad (9)$$

where m is a hyper-parameter which belong to convex space $\mathbb{L}$:

$$\mathbb{L} = \left\{ \ell \in \mathbb{R}^d : \ell_i \in [-c_i, c_i], c_i > 0, i = \{1, ..., d\} \right\}$$

[6] Alessio BENAVOLI et Marco ZAFFALON. “Prior near ignorance for inferences in the k-parameter exponential family”. In : *Statistics* 49.5 (2015), p. 1104-1140

Remark

$\mathcal{M}_0^{\mu}$ satisfy the four minimal properties that model of prior ignorance require : invariance, near-ignorance, learning and convergence (more details [6]).

Imprecise Linear Discriminant Analysis

By applying Baye's rule (8) (or equation [6, eq 17]), we get a set of posterior distribution :

$$\mathscr{M}_{n_k}^{\mu_k} = \left\{ \mu_k \middle| \overline{\boldsymbol{x}}_{n_k}, m \propto \mathcal{N}\left(\frac{\ell + n_k \overline{\boldsymbol{x}}_{n_k}}{n_k}, \frac{1}{n_k}\widehat{\Sigma} \right), \right\} \quad (10)$$

where $\overline{\boldsymbol{x}}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} x_{i,k}$ and $\ell \in \mathbb{L}$, and :

$$\inf_{\mathscr{M}_{n_k}^{\mu_k}} \mathbb{E}[\boldsymbol{\mu}_k \mid \overline{\boldsymbol{x}}_{n_k}, \ell] = \underline{\mathbb{E}}[\boldsymbol{\mu}_k \mid \overline{\boldsymbol{x}}_{n_k}, m] = \frac{-\ell + n_k \overline{\boldsymbol{x}}_{n_k}}{n} \quad (11)$$

$$\sup_{\mathscr{M}_{n_k}^{\mu_k}} \mathbb{E}[\boldsymbol{\mu}_k \mid \overline{\boldsymbol{x}}_{n_k}, \ell] = \overline{\mathbb{E}}[\boldsymbol{\mu}_k \mid \overline{\boldsymbol{x}}_{n_k}, m] = \frac{\ell + n_k \overline{\boldsymbol{x}}_{n_k}}{n_k} \quad (12)$$

Imprecise Linear Discriminant Analysis

The two last estimations describe a convex set around μ :

$$\mathbb{G}_k = \left\{ \hat{\mu}_k \in \mathbb{R}^d \middle| \begin{array}{l} \hat{\mu}_{i,k} \in \left[\frac{-c_i + n_k \overline{\boldsymbol{x}}_{i,n_k}}{n_k}, \frac{c_i + n_k \overline{\boldsymbol{x}}_{i,n_k}}{n_k} \right], \\ \forall i = \{1, ..., d\} \end{array} \right\}$$

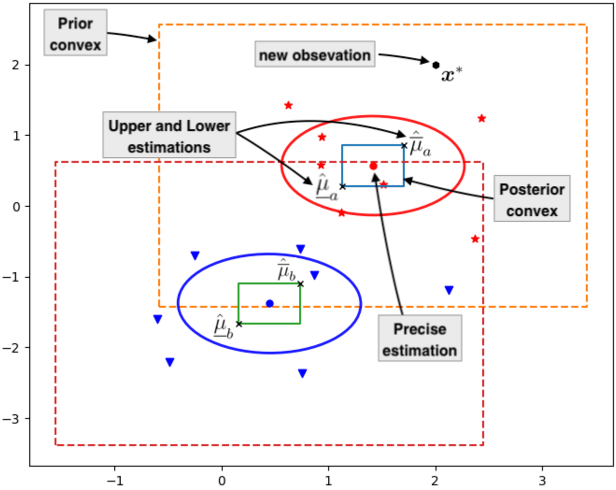
That we use as constraint in on our two optimisation problems.

$$\underline{P}(\boldsymbol{x}^* | y = m_a) \iff \underline{\hat{\mu}}_a = \arg\max_{\hat{\mu}_a \in \mathbb{G}_a} \frac{1}{2} \hat{\mu}_a^T \hat{\Sigma}^{-1} \hat{\mu}_a + \boldsymbol{x}^{*T} \hat{\Sigma}^{-1} \hat{\mu}_a \quad \text{(NPQB)}$$

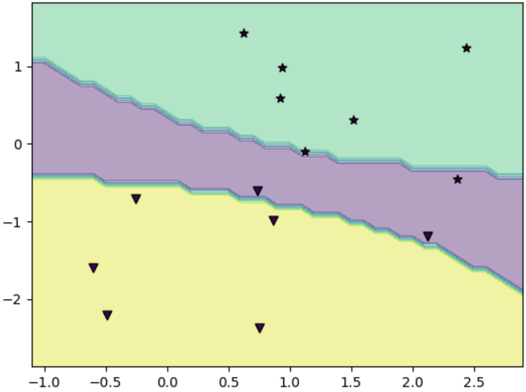
$$\overline{P}(\boldsymbol{x}^* | y = m_b) \iff \hat{\overline{\mu}}_b = \arg\min_{\hat{\mu}_b \in \mathbb{G}_b} \frac{1}{2} \hat{\mu}_b^T \hat{\Sigma}^{-1} \hat{\mu}_b + \boldsymbol{x}^{*T} \hat{\Sigma}^{-1} \hat{\mu}_b \quad \text{(PQB)}$$

First problem non-convex → solved through B&B method.

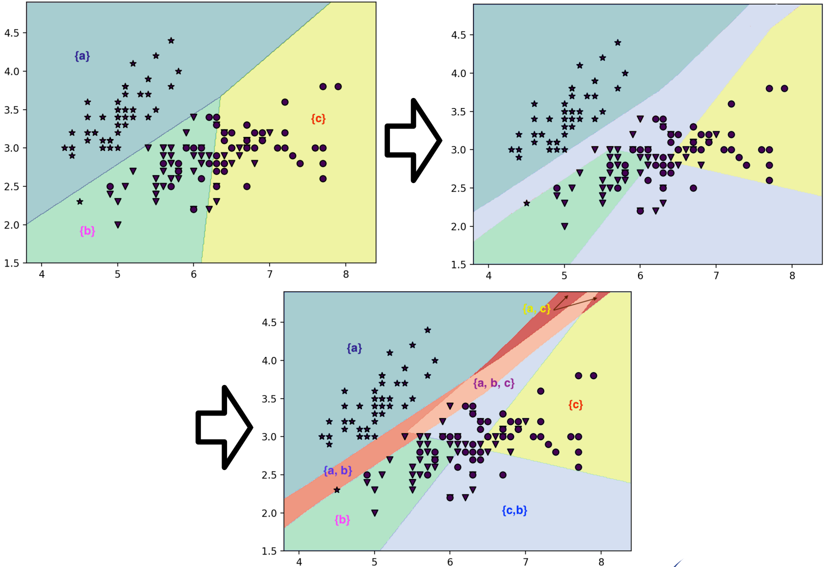
Example


Prior convex
new obsevation
x^*
Upper and Lower estimations
$\hat{\overline{\mu}}_a$
$\hat{\underline{\mu}}_a$
Posterior convex
$\hat{\overline{\mu}}_b$
$\hat{\underline{\mu}}_b$
Precise estimation
2
1
0
−1
−2
−3
−1
0
1
2
3

Example (cont..)



Another Example with 3 class



Experiments

Average utility-discounted accuracy measure of [4]

$$u(y, \widehat{Y}_M) = \begin{cases} 0 & \text{if } y \notin \widehat{Y}_M \\ \frac{\alpha}{|\widehat{Y}_M|} - \frac{\beta}{|\widehat{Y}_M|} & \text{else} \end{cases}$$

Where u_{65} with $(\alpha, \beta) = (1.6, 0.6)$ and u_{80} with $(\alpha, \beta) = (2.2, 1.2)$.

#	Name	# Obs.	# Regr.	# Classes
a	iris	150	4	3
b	seeds	210	7	3
c	glass	214	9	6

#	DLA	IDLA		Inference time
		u_{65}	u_{80}	
a	0.961	0.969	0.975	0.56 sec.
b	0.959	0.959	0.962	1.50 sec.
c	0.594	0.589	0.642	8.66 sec

Overview

- Classification
  - Decision Making
  - Discriminant Analysis
- Imprecise Classification
  - Imprecise Decision
  - Imprecise Linear discriminant analysis
- **Future work**
- Conclusions

heudiasyc

Imprecise Quadratic discriminant analysis

(1) Release homoscedasticity assumption, i.e. $\Sigma_k \neq \Sigma$

$$\underline{\hat{\mu}}_a = \arg\max \frac{1}{2}\hat{\mu}_a^T \hat{\Sigma}_k^{-1} \hat{\mu}_a - \boldsymbol{x}^{*T} \hat{\Sigma}_k^{-1} \hat{\mu}_a$$

$$s.t. \frac{-c_j + n\overline{x}_{j,n}}{n} \leq \hat{\mu}_{j,a} \leq \frac{c_i + n\overline{x}_{j,n}}{n} \qquad \text{(PQB)}$$

$$\forall j = \{1, ..., d\}$$

- *Making imprecise* $P(y = m_a) = [\underline{P}(y = m_a), \overline{P}(y = m_a)]$ *and to solve :*

$$\inf_{P_{X|y} \in \mathscr{P}_1, P_y \in \mathscr{P}_2} \frac{P(\boldsymbol{x}^* | y = m_a) P(y = m_a)}{P(\boldsymbol{x}^* | y = m_b) P(y = m_b)} > 1$$

Imprecise Quadratic discriminant analysis

Space Convex Matrices $\mathbf{S}_+$

(2) Make imprecise the covariance matrice (i.e. Σ_k or Σ) by using a prior Wishart distribution :

$$\underline{\Sigma}_k = \inf_{\Omega \in \mathbf{S}_+^n} \mathbb{E}[\Sigma_k | X, y = m_k, \tau_0, \Omega] \tag{13}$$

$$\underline{\Sigma}_k = \inf_{\Omega \in \mathbf{S}_+^n} \frac{\Omega + (n-1)\widehat{\Sigma}_k^{\mathrm{MLE}}}{n + \tau_0} \tag{14}$$

where $\widehat{\Sigma}_k^{\mathrm{MLE}}$ is the maximal likelihood estimator of covariance matrice Σ_k and $\mathbf{S}_+^n$ is a convex space of families of positive semi-definite positive matrices.

Imprecise Quadratic discriminant analysis

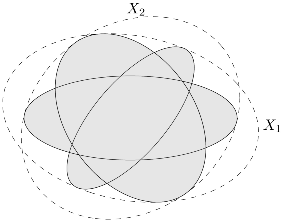
Space Convex Matrices $\mathbf{S}_+$

In [2], we can find a good intuitions for minimize the last optimization problem, where Φ_ϵ is a perturbation in the neighbourhood of Ω_0 prior parameter value, and $||\cdot||_F$ is Frobenius norm.

$$\arg\min_{\Omega_0 \in \mathbf{S}_+^n} \underline{\Sigma} = \frac{\Omega_0 + (n-1)\widehat{\Sigma}_e}{n+\tau_0}$$

$$\text{s.t.} \quad \underline{\Sigma} \succeq X_i, \quad \forall X_i \in \mathbf{S}_+^n, i = \{1, ..., m\}$$

$$\mathbf{S}_+^n = \{\Omega_0 \mid ||\Omega_0 - \Phi_\epsilon||_F \preceq \Omega_0 \preceq ||\Omega_0 + \Phi_\epsilon||_F\}$$



Imprecise Quadratic discriminant analysis
Space Convex of eigenvalues or eigenvectors

(3) Imprecise eigenvalues and eigenvectors of Σ_k.

We'll propose to use $\widehat{\Omega}$ estimation of [3, §3], i.e $\widehat{\Omega} = \frac{tr(\Sigma_k^{MLE})}{d}$, and then applying it the spectral decomposition :

$$\frac{\widehat{\Omega} + (n-1)\widehat{\Sigma}_k^{\text{MLE}}}{n+\tau_0} \iff \frac{tr(\sum_{j=1}^{d} \lambda_j u_j u_j^t)}{d(n+\tau_0)}\mathbb{I} + \frac{n-1}{n+\tau_0}\sum_{j=1}^{d} \lambda_j u_j u_j^t \quad (15)$$

$$\sum_{j=1}^{d} \lambda_j \left[\frac{tr(u_j u_j^t)}{d}\mathbb{I} + (n-1) u_j u_j^t \right] \quad (16)$$

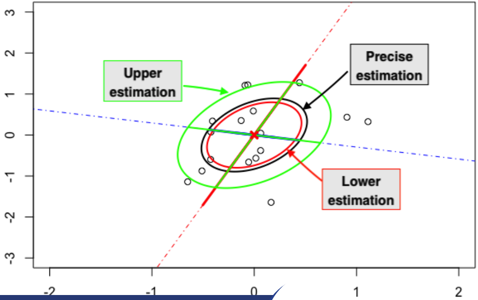
Imprecise Quadratic discriminant analysis

Space Convex of eigenvalues or eigenvectors

In [3], it has been proven that eigenvalues have estimations either biased high (overestimated) or biased low (underestimated) for small and noisy samples.
Then, we could assume the variability of direction is "correctly" estimated (i.e eigenvectors)

$$\overline{\lambda} = \arg\max_{\lambda \in \mathbf{S}_+^n} \sum_{j=1}^{d} \lambda_j \left[\frac{tr(u_j u_j^t)}{d} \mathbb{I} + (n-1) u_j u_j^t \right]$$

$$s.t \;\; \mathbf{S}_+^n = \left\{ \widehat{\Sigma} = \sum_{j=1}^{d} \lambda_j v v^t \;\middle|\; \underline{\Sigma} \preceq \widehat{\Sigma} \preceq \overline{\Sigma} \right\}$$



Overview

- Classification
  - Decision Making
  - Discriminant Analysis
- Imprecise Classification
  - Imprecise Decision
  - Imprecise Linear discriminant analysis
- Future work
- Conclusions

Conclusions

Imprecise Analysis Discriminant Classification

- Increasing in imprecision on the estimators has allowed us to be more cautious in doubt and to improve the prediction of classification [7].
- More experiments with all imprecise components.
- Creation of new imprecise statistic models for a sensibility analysis and a more (cautious) robust prediction.

References

Jerome FRIEDMAN, Trevor HASTIE et Robert TIBSHIRANI. *The elements of statistical learning*. T. 1. Springer series in statistics New York, 2001.

Stephen BOYD et Lieven VANDENBERGHE. *Convex optimization*. Cambridge university press, 2004.

Santosh SRIVASTAVA. *Bayesian Minimum Expected Risk Estimation of Distributions for Statistical Learning*. University of Washington, 2007.

Marco ZAFFALON, Giorgio CORANI et Denis MAUÁ. "Evaluating credal classifiers by utility-discounted predictive accuracy". In : *International Journal of Approximate Reasoning* 53.8 (2012), p. 1282-1301.

James O BERGER. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 2013.

Alessio BENAVOLI et Marco ZAFFALON. "Prior near ignorance for inferences in the k-parameter exponential family". In : *Statistics* 49.5 (2015), p. 1104-1140.

Yonatan-Carlos CARRANZA-ALARCON et Sébastien DESTERCKE. "Analyse Discriminante Imprécise basé sur l'inference Bayésienne robuste". In : *27 emes rencontres francophones sur la logique floue et ses applications* (2018).


Thank
You!


ANY QUESTIONS?
PLEASE NO HARD QUESTIONS
memegenerator.net