# KRIGGING

: SOURCE: PROBABILISTIC INTERVAL PREDICTOR BASED ON DISSIMILARITY FUNCTIONS

$$Z = [z_1 \dots z_N] = \begin{bmatrix} z_{1,1} & \cdots & z_{N,1} \\ \vdots & \ddots & \vdots \\ z_{1,c} & \cdots & z_{N,c} \end{bmatrix}$$

$$z_i = \begin{bmatrix} z_{i,1} \\ \vdots \\ z_{i,c} \end{bmatrix}$$

SYSTEM OUTPUT; UNKNOWN SYSTEM; PARAMETRIC UNCERTAINTY NOISE

$$y_{k} = f_0(x_k, w_k)$$

$k$: TIME INSTANT; $x_k$: PAST INPUTS/OUTPUTS

DISSIMILARITY FUNCTION; GOAL IS TO CHECK HOW DISSIMILAR $z$ IS TO $D$; $D = \{z_1 \dots z_N\}$ MEASURES; USUALLY $w_i = 1 \ \forall i$

$$J_\gamma(z, D) = \min_{\lambda_1 \dots \lambda_N} \sum_{i=1}^{N} w_i \lambda_i^2 + \gamma \sum_{i=1}^{N} |\lambda_i|$$

$\gamma \geq 0$

$$\text{s.t. } z = \sum_{i=1}^{N} \lambda_i z_i$$

$$1 = \sum_{i=1}^{N} \lambda_i$$

$Z = [z_1 \dots z_N]$; $\bar{z} = N^{-1} Z u$; $u = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}$; $u_i = 1 \ \forall i$; $\bar{z} = \begin{bmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_c \end{bmatrix}$

$$J_0(z, D) = N^{-1} + (z - \bar{z})^T (Z Z^T - N \bar{z} \bar{z}^T)^{-1} (z - \bar{z})$$

$J_0$: $J_\gamma, \gamma = 0$

GOAL: COMPARE WITH $J$ FOR QDA//

FOR CLASSIFICATION: SELECT $i$ S.T. $J_\gamma(z, D_i) \leq J_\gamma(z, D_k) \ \forall k$

# QDA

: SOURCE: SCIKIT DOCUMENTATION

FOR CLASSIFICATION: SELECT $i$ S.T. $\log P(y = i \mid x) \geq \log P(y = k \mid x) \ \forall k$

$$\log P(y = k \mid x) = \frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (z - \bar{z})^T \Sigma_k^{-1} (z - \bar{z}) + \log P(y = k) + C_{ST}$$

$$\Sigma_k = \begin{bmatrix} VAR(z_{i,1}) & COV(z_{i,2}, z_{i,1}) & \cdots & COV(z_{i,c}, z_{i,1}) \\ COV(z_{i,1}, z_{i,2}) & VAR(z_{i,2}) & \cdots & COV(z_{i,c}, z_{i,2}) \\ \vdots & \vdots & \ddots & \vdots \\ COV(z_{i,1}, z_{i,c}) & COV(z_{i,2}, z_{i,c}) & \cdots & VAR(z_{i,c}) \end{bmatrix}$$

$$COV(A, B) = \frac{\sum_{i=1}^{n} (A_i - \mu_A)(B_i - \mu_B)}{n}$$

$$COV(A, B) = COV(B, A)$$

$$VAR(A) = COV(A, A)$$

$$\Sigma_k = \frac{1}{N} (Z - \bar{z} u^T)(Z - \bar{z} u^T)^T = N^{-1} [Z - \bar{z} u^T] (Z^T - u \bar{z}^T)$$

$$\Sigma_k = N^{-1} (Z Z^T - Z u \bar{z}^T - \bar{z} u^T Z^T + \bar{z} u^T u \bar{z}^T)$$

$$\Sigma_k = N^{-1} Z Z^T - N^{-1} Z u \bar{z}^T - (N^{-1} Z u \bar{z}^T)^T + N^{-1} \bar{z} u^T u \bar{z}^T$$

$\bar{z} = N^{-1} Z u$

$$\bar{z} u^T u \bar{z}^T = \begin{bmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_N \end{bmatrix} [1 \dots 1] \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} [\bar{z}_1 \dots \bar{z}_N] = N \bar{z} \bar{z}^T$$

$u^T u = N$

$$\Sigma_k = N^{-1} Z Z^T - \bar{z} \bar{z}^T - \cancel{(\bar{z} \bar{z}^T)^T} + \cancel{\bar{z} \bar{z}^T}$$

$$\Sigma_k = N^{-1} Z Z^T - \bar{z} \bar{z}^T$$

CONSTANTS HAVE NO COMPARATIVE INTEREST

$$-J = \frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (z - \bar{z}_k)^T \Sigma_k (z - \bar{z}_k) + \log P(y = k) + \cancel{C_{ST}}$$

$$+J = -\frac{1}{2} \log |N^{-1} Z Z^T - \bar{z} \bar{z}^T| + \frac{1}{2} (z - \bar{z})^T (N^{-1} Z Z^T - \bar{z} \bar{z}^T)^{-1} (z - \bar{z}) \mp \log P(y = k)$$

# COMPARISON:

$$J_{0,KRIGGING}(z,D) = N^{-1} + \underbrace{(z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z)}$$

$$J_{QDA} = \frac{-1}{2}\log|N^{-1}ZZ^T - \bar z\bar z^T| + \underbrace{\frac{1}{2}(z-\bar z)^T(N^{-1}ZZ^T - \bar z\bar z^T)^{-1}(z-\bar z)} - \log P(y=K)$$

$$J_{QDA} = \frac{-1}{2}\log|N^{-1}ZZ^T - \bar z\bar z^T| + \frac{1}{2}(z-\bar z)^T(N^{-1}ZZ^T - \bar z\bar z^T)^{-1}(z-\bar z) - \log P(y=K)$$

$$J_{QDA} = \frac{-1}{2}\log|N^{-1}ZZ^T - \bar z\bar z^T| + \frac{1}{2}(z-\bar z)^T(N^{-1}(ZZ^T - N\bar z\bar z^T))^{-1}(z-\bar z) - \log P(y=K)$$

$$J_{QDA}\cdot 2 = N\underbrace{(z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z)} - \underbrace{\log|\overbrace{N^{-1}ZZ^T - \bar z\bar z^T}^{\Sigma_K}| - 2\log P(y=K)}$$

IF THE TRAINING SET IS BALANCED, N BECOMES A CONSTANT

CONSTANT WITH NO COMPARATIVE VALUE

$$\frac{J_{QDA}\cdot 2}{N} = (z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z) - \cancel{\frac{2}{N_T}} - \frac{\log|N^{-1}ZZ^T - \bar z\bar z^T|}{N}$$

$$J_{\cdot QDA} = \frac{J_{QDA}\cdot 2}{N} + \frac{2}{N_T} = \underbrace{(z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z)}_{\text{SAME TERM}} - \underbrace{N^{-1}\log|N^{-1}ZZ^T - \bar z\bar z^T|}$$

TERM $N^{-1}$ IS SCALED BY $\log|\Sigma_K|$

SIGN IS OPPOSITE TO KRIGGING

## LDA: QDA, BUT $\Sigma_K = \Sigma \ \forall K$

CONSTANT

$$2J_{LDA} = N(z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z) - \frac{2N}{N_T} - \boxed{\log|N^{-1}ZZ^T - \bar z\bar z^T|}$$

$$2J_{LDA} + \log|\Sigma| = N\underbrace{(z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z)}_{\text{SAME TERM}} - \frac{2N}{N_T}$$

BUT SCALED BY N

AS WITH $+N^{-1}$, PENALIZES SMALLER N BUT THIS ONE ALSO CONSIDERS $N_T$

HIGHER N MAY HELP OR PENALIZE, DEPENDING ON $\Sigma$ AND $z-\bar z$

IF TRAINING SET IS BALANCED

$$J_{0,KRIGGING} = \overbrace{N^{-1}}^{\text{CONSTANT}} + (z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z) \rightarrow N\, J_{0,KRIGGING} - 1 = N(z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z)$$

$$2J_{LDA} + \log|\Sigma| + \frac{2N}{N_T} = N(z-\bar z)^T(ZZ^T - N\bar z\bar z^T)^{-1}(z-\bar z)$$

KRIGGING WITH $\gamma=0$ IS EQUIVALENT TO LDA, AS LONG AS CLASSES ARE BALANCED//

REASON: COST FUNCTIONS VALUES ARE THE SAME, ASIDE OF LINEAR COMBINATIONS THAT DON'T AFFECT ORDER.

$\{ J_{\gamma,KRIGGING} \}$?

# KRIGGING $\gamma \neq 0$ :

$$J_\gamma(z, D) = \min_{\lambda_1 \dots \lambda_N} \sum_{i=1}^{N} \lambda_i^2 + \gamma \sum_{i=1}^{N} |\lambda_i|$$

$$\text{s.t. } \quad z = \sum_{i=1}^{N} \lambda_i z_i$$

$$1 = \sum_{i=1}^{N} \lambda_i$$

$$\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{bmatrix}$$

$$\left.\begin{matrix} u^T \lambda = 1 \\ Z\lambda = z \end{matrix}\right\} \begin{bmatrix} Z \\ u^T \end{bmatrix} \lambda = \begin{bmatrix} z \\ 1 \end{bmatrix} \to \lambda = \begin{bmatrix} Z \\ u^T \end{bmatrix}^{-1} \begin{bmatrix} z \\ 1 \end{bmatrix}$$

IT WON'T PROBABLY WORK
(N>C)