Complément de cours

Pour le GLRT, voir les notes déposées sur Moodle.

Pour tester $\mathcal{H}_0 : Y \sim p_0$ contre $\mathcal{H}_1 : Y \sim p_1$, la fonction critique du lemme de Neyman-Pearson (NP) est :

$$\varphi_{NP}(y) = \begin{cases} 1 & \text{si } \Lambda(y) > \lambda \\ \gamma & \text{si } \Lambda(y) = \lambda \\ 0 & \text{si } \Lambda(y) < \lambda \end{cases}$$

avec $\Lambda = p_1 / p_0$ et où λ et γ sont calculés de manière à ce que

$$\beta_{\varphi_{NP}}(\Theta_0) = \alpha .$$

La puissance du test NP est alors la fonction de puissance en Θ_1, c'est-à-dire la valeur $\beta_{\varphi_{NP}}(\Theta_1)$

On remarque que φ_{NP} s'écrit encore

$$\varphi_{NP} = \mathbb{1}_{[\Lambda > \lambda]} + \gamma \mathbb{1}_{[\Lambda = \lambda]}$$

Calculons alors $\beta_{\varphi_{NP}}(\Theta_0)$

Après quelques calculs (cf. notes de cours 2017), on obtient

$$\beta_{\varphi_{NP}}(\theta_0) = \mathbb{P}[\Lambda(Y) > \lambda] + \gamma\,\mathbb{P}[\Lambda(Y) = \lambda].$$

On doit donc résoudre l'équation :

$$\mathbb{P}[\Lambda(Y) > \lambda] + \gamma\,\mathbb{P}[\Lambda(Y) = \lambda] = \alpha$$

Dans les nombreux cas pratiques où $\mathbb{P}[\Lambda(Y) = \lambda] = 0$, on peut choisir γ quelconque et on calcule λ en résolvant l'équation $\underline{\mathbb{P}[\Lambda(Y) > \lambda] = \alpha}$. On choisira généralement $\gamma = 0$ de sorte que φ_{NP} devient :

$$\varphi_{NP}(y) = \begin{cases} 1 \text{ si } \Lambda(y) > \lambda \\ 0 \text{ si } \Lambda(y) \leq \lambda \end{cases}$$

Autrement dit, en posant $\Gamma = [\Lambda > \lambda]$

$$= \{y \in \mathbb{R}^d : \Lambda(y) > \lambda\}$$

$\varphi_{NP} = \mathbb{1}_\Gamma$. Comparer avec le test T_{MPE} ! C'est la même forme de fonction critique avec un seuil différent !

La décision $\mathcal{D}_{NP}$ est alors $\mathcal{D}_{NP} = \varphi_{NP}(Y)$

$$= \mathbb{1}_\Gamma(Y)$$

Nous savons que $\beta_{\varphi_{NP}}(0) = \alpha$. Calculons la puissance $\beta_{\varphi_{NP}}(\theta_1)$ du test de Neyman-Pearson

En effet, dans la littérature sur le sujet, on emploie le terme de test pour désigner la fonction critique, comme on le fait pour le test MPE

Nous faisons le calcul dans le cas où $\mathbb{P}[\Lambda = \lambda] = 0$. On a :

$$\beta_{\varphi_{NP}}(\theta_1) = \int \varphi_{NP}(y)\, p_1(y)\, dy$$

$$= \int \mathbb{1}_\Gamma(y)\, p_1(y)\, dy$$

$$= \int_\Gamma p_1(y)\, dy$$

$$\Rightarrow \beta_{\varphi_{NP}}(\theta_1) = \mathbb{P}[Y_1 \in \Gamma] \text{ pour tout } Y_1 \sim p_1$$

On retrouve bien que la puissance $\beta_{\varphi_{NP}}(\theta_1)$ du test NP est la probabilité d'accepter $\mathcal{H}_1$ lorsque cette hypothèse est vraie.

Détection d'un signal déterministe dans un bruit blanc Gaussien : comparaison entre le test de Neyman-Pearson et le test GLRT

o ——————— o◯o ——————— o

On suppose que $\begin{cases} \mathcal{H}_0 : Y \sim \mathcal{N}(0, \sigma^2 I_N) \\ \mathcal{H}_1 : Y \sim \mathcal{N}(A\xi_0, \sigma^2 I_N) \end{cases}$

Avant de commencer le calcul, on remarque que si A est connu, le problème est bien un test d'hypothèses simples. Avec les notations du cours :

- p_0 est la distribution $\mathcal{N}(0, \sigma^2 I_N)$
- p_1 est la distribution $\mathcal{N}(A\xi_0, \sigma^2 I_N)$
- $\mathcal{F} = \{p_0, p_1\}$, $\Theta_0 = 0$, $\Theta_1 = 1$. Le paramètre est simplement l'indice de la densité
- $\mathcal{X} = \{0, 1\}$, $\mathcal{X}_0 = \{0\}$, $\mathcal{X}_1 = \{1\}$

Le test clairvoyant est le test de Neyman-Pearson calculé en supposant connu A.

On suppose que $\begin{cases} \mathcal{H}_0 : Y \sim \mathcal{N}(0, \sigma^2 I_N) \\ \mathcal{H}_1 : Y \sim \mathcal{N}(A\theta_0, \sigma^2 I_N) \end{cases}$

1) Calcul du test clairvoyant

$\forall\, y \in \mathbb{R}^N$, on a $\Lambda(y) = \dfrac{p_1(y)}{p_0(y)}$ avec :

$$p_0(y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}\|y\|_2^2\right)$$ ($\|\cdot\|_2$: norme euclidienne dans $\mathbb{R}^N$)

$$p_1(y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}\|y - A\xi_0\|_2^2\right)$$

$$\Lambda(y) = \exp\left(-\frac{1}{2\sigma^2}\left(\|y - A\xi_0\|_2^2 - \|y\|_2^2\right)\right)$$

$$\Rightarrow \operatorname{Log}\Lambda(y) = -\frac{1}{2\sigma^2}\left(\|y - A\xi_0\|_2^2 - \|y\|_2^2\right)$$

$$= -\frac{1}{2\sigma^2}\left(\cancel{\|y\|_2^2} - 2Ay^T\xi_0 + A^2\|\xi_0\|_2^2 - \cancel{\|y\|_2^2}\right)$$

$$\log \Lambda(y) = \frac{A}{\sigma^2} y^T \xi_0 - \frac{A^2}{2\sigma^2} \quad \text{car } \|\xi_0\|_2 = 1.$$

Pour $A > 0$, le test de Neyman - Pearson est donc :

$$\forall y \in \mathbb{R}^N, \quad \varphi^+_{NP}(y) = \begin{cases} 1 & \text{si } y^T \xi_0 > \lambda_+ \\ 0 & \text{sinon} \end{cases}$$

On détermine λ en résolvant l'équation :

$$\alpha^+_{NP} = \mathbb{P}\left[Y_0^T \xi_0 > \lambda_+ \right] = \alpha \quad \text{où } Y_0 \sim \mathcal{N}(0, \sigma^2 I_N)$$

On a $\mathbb{E}\left[Y_0^T \xi_0\right] = 0$.

$$\operatorname{var}\left\{ Y_0^T \xi_0 \right\} = \mathbb{E}\left[\xi_0^T Y_0 Y_0^T \xi_0 \right] = \xi_0^T \sigma^2 I_N \xi_0 = \sigma^2 \|\xi_0\|^2 = \sigma^2$$

Donc $\boxed{Y_0^T \xi_0 \sim \mathcal{N}(0, \sigma^2)}$

$$\alpha^+_{NP} = \mathbb{P}\left[Y_0^T \xi_0 > \lambda_+ \right] = 1 - \mathbb{P}\left[Y_0^T \xi_0 \leq \lambda_+ \right]$$

$$= 1 - \mathbb{P}\left[\frac{Y_0^T \xi_0}{\sigma} \leq \frac{\lambda_+}{\sigma} \right] = Q\left(\lambda_+ / \sigma\right)$$

Donc : $\boxed{\lambda_+ = \sigma Q^{-1}(\gamma)}$

La puissance de φ^+_{NP} est :

$$\beta^+_{NP} = \mathbb{P}\left[Y_1^T \xi_0 > \lambda_+ \right] \quad \text{où } Y_1 \sim \mathcal{N}(A\xi_0, \sigma^2 I_N)$$

$$\mathbb{E}\left[Y_1^T \xi_0\right] = \mathbb{E}\left[Y_1^T\right]\xi_0 = A\,\xi_0^T \xi_0 = A\,\|\xi_0\|_2^2 = A$$

$$\operatorname{var}\left\{Y_1^T \xi_0\right\} = \mathbb{E}\left[\left(Y_1^T\xi_0 - \mathbb{E}\left[Y_1^T\xi_0\right]\right)^2\right]$$

$$= \mathbb{E}\left[\left(\left(Y_1^T - \mathbb{E}\left[Y_1^T\right]\right)\xi_0\right)^2\right]$$

$$= \mathbb{E}\left[\left(Z^T\xi_0\right)^2\right] \quad \text{avec } Z = Y_1 - \mathbb{E}\left[Y_1\right]$$

$Z \sim \mathcal{N}(0, \sigma^2 I_N) \Rightarrow Z^T\xi_0 \sim \mathcal{N}(0, \sigma^2)$ par le même calcul que précédemment.

On a donc $\operatorname{var}\left\{Y_1^T\theta_0\right\} = \sigma^2$. On obtient donc que :

$Y_1^T\theta_0 \sim \mathcal{N}(A, \sigma^2)$. Il s'ensuit que :

$$\beta_{NP}^+ = \mathbb{P}\left[Y_1^T\xi_0 > \lambda_+\right] = \mathbb{P}\left[\frac{Y_1^T\xi_0 - A}{\sigma} > \frac{\lambda_+ - A}{\sigma}\right]$$

$$= 1 - \mathbb{P}\left[\frac{Y_1^T\xi_0 - A}{\sigma} \leq \frac{\lambda_+ - A}{\sigma}\right] = Q\left(\frac{\lambda_+ - A}{\sigma}\right)$$

Comme $\frac{\lambda_+}{\sigma} = Q^{-1}(\gamma)$, $\boxed{\beta_{NP}^+ = Q\left(Q^{-1}(\gamma) - \frac{A}{\sigma}\right)}$

Lorsque $\frac{A}{\sigma} \to \infty$, $\beta_{NP}^+ \to 1$ et lorsque $\frac{A}{\sigma} \to 0$, $\beta_{NP}^+ \to \gamma$

$Q^{-1}(\gamma) - \frac{A}{\sigma} < Q^{-1}(\gamma) \Rightarrow \beta_{NP}^+ \geq \gamma$. Le test est dit non biaisé.

Pour $A < 0$, le test de Neyman - Pearson est donc :

$$\forall y \in \mathbb{R}^N, \quad \varphi^-_{NP}(y) = \begin{cases} 1 & \text{si } y^T \theta_0 < \lambda_- \\ 0 & \text{sinon} \end{cases}$$

On détermine λ_- en résolvant l'équation :

$$\alpha^-_{NP} = \mathbb{P}\left[Y_0^T \xi_0 < \lambda_- \right] = \alpha \quad \text{où } Y_0 \sim \mathcal{N}(0, \sigma^2 I_N)$$

Comme précédemment $Y_0^T \xi_0 \sim \mathcal{N}(0, \sigma^2)$.

Donc $\quad \alpha^-_{NP} = \mathbb{P}\left[Y_0^T \xi_0 \leq \lambda_- \right] = Q(-\lambda_- / \sigma)$

On en déduit $\boxed{\lambda_- = -\sigma Q^{-1}(\gamma)}$

$$\beta^-_{NP} = \mathbb{P}\left[Y_1^T \xi_0 \leq \lambda_- \right] \quad \text{avec } Y_1 \sim \mathcal{N}(A\theta_0, \sigma^2 I_N)$$

Comme précédemment $Y_1^T \xi_0 \sim \mathcal{N}(A, \sigma^2)$. Aussi :

$$\beta^-_{NP} = \mathbb{P}\left[\frac{Y_1^T \xi_0 - A}{\sigma} \leq \frac{\lambda_- - A}{\sigma} \right]$$

$$= Q\left(-\frac{\lambda_- - A}{\sigma} \right)$$

$$-\frac{\lambda_-}{\sigma} = Q^{-1}(\gamma) \Rightarrow \boxed{\beta^-_{NP} = Q\left(Q^{-1}(\gamma) + \frac{A}{\sigma} \right)}$$

Pour $A > 0$ ou $A < 0$, la puissance de φ_{NP} est la même et vaut :

$$\boxed{\beta_{NP}^{-} = Q\left(Q^{-1}(\gamma) - \frac{|A|}{\sigma}\right)}$$

Le test clairvoyant est donc le test φ_{NP}^{+} si on savait que $A > 0$ et le test φ_{NP}^{-} pour $A < 0$. Ce test ne s'applique pas si A est inconnu mais sert de référence pour mesurer la perte de performance du GLRT.

Ce test clairvoyant, qui connaît A et donc son signe, est :

$$\text{Si } A > 0, \quad \varphi^*(y) = \begin{cases} 1 & \text{si } y^T\theta_0 > \sigma Q^{-1}(\gamma) \\ 0 & \text{sinon} \end{cases}$$

$$\text{Si } A < 0, \quad \varphi^*(y) = \begin{cases} 1 & \text{si } y^T\theta_0 < \lambda_- \\ 0 & \text{sinon} \end{cases}$$

La puissance de ce test clairvoyant est :

$$\boxed{\beta^* = Q\left(Q^{-1}(\gamma) - \frac{|A|}{\sigma}\right)}$$

Il n'y a donc pas de test UMP puisque, suivant le signe de A, nous avons des tests "most powerful" différents !

2) Le test GLRT est la fonction critique

$$\varphi_{GLRT}(y) = \begin{cases} 1 & \text{si } \Lambda_G(y) > \lambda_G \\ 0 & \text{sinon} \end{cases}$$

où $\Lambda_G(y) = \dfrac{\sup_A p_1(y;A)}{p_0(y)}$;

$$p_1(y;A) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\|y - A\xi_0\|_2^2\right)$$

λ_G est calculé de manière à ce que la taille $\alpha_{GLRT} = \beta_{\varphi_{GLRT}}(0)$ soit égale à α.

On a $\sup_A p_1(y;A) = p_1(y;\hat{A}_{MLE}(y))$

avec $\hat{A}_{MLE}(y)$ est la valeur qui maximise $p_1(y;A)$:

$$\hat{A}_{MLE}(y) = \underset{A}{\mathrm{argmax}}\ p_1(y;A)$$

MLE = Maximum Likelihood Estimate.

Compte-tenu de la fonction $p_1(y;A)$ on obtient $\hat{A}_{MLE}$ comme solution de l'équation $\frac{\partial p_1(y;A)}{\partial A} = 0$.

On remarque ici que chercher un maximum de $p_1(y;A)$ équivaut à chercher un maximum de $\Lambda(y)$ puisque p_0 ne dépend pas de A !

Comme le log croît, on a donc :

$$\hat{A}_{MLE} = \underset{A}{\mathrm{argmax}}\ \log\Lambda(y;A)$$

avec $\Lambda(y;A) = p_1(y;A)/p_0(y)$

$$\log \Lambda(y;A) = \frac{A}{\sigma^2} y^T \xi_0 - \frac{A^2}{2\sigma^2}$$

$$\frac{\partial \log \Lambda(y;A)}{\partial A} = \frac{1}{\sigma^2} y^T \xi_0 - \frac{A}{\sigma^2}$$

$$\frac{\partial \log \Lambda(y;A)}{\partial A} = 0 \Leftrightarrow \boxed{\hat{A}_{MLE}(y) = y^T \xi_0}$$

$p_1(y;A) = K \exp\left(-\frac{1}{2\sigma^2} \|y - A\xi_0\|_2^2\right)$ est la densité de probabilité de l'observation sous $\mathcal{H}_1$, avec $K = 1/(2\pi\sigma^2)^{N/2}$.

La densité de l'observation sous $\mathcal{H}_0$ ne dépend pas de A. Le rapport de vraisemblance généralisé est donc :

$$\forall y \in \mathbb{R}^N, \quad \Lambda_G(y) = \frac{\sup_A p_1(y;A)}{p_0(y)}$$

$$= \frac{p_1(y; \hat{A}_{MLE}(y))}{p_0(y)}$$

On a donc :

$$\Lambda_G(y) = \exp\left(-\frac{1}{2\sigma^2}\left(\|y - \hat{A}_{MLE}(y)\xi_0\|_2^2 - \|y\|_2^2\right)\right)$$

$$\Rightarrow \log \Lambda_G(y) = -\frac{1}{2\sigma^2}\left(\|y - \hat{A}_{MLE}(y)\xi_0\|_2^2 - \|y\|_2^2\right)$$

$$= \frac{\hat{A}_{MLE}(y)}{\sigma^2} y^T \xi_0 - \frac{1}{2\sigma^2} \hat{A}_{MLE}^2(y)$$

On obtient :

$$\mathrm{Log}\, \Lambda_G(y) = \frac{(y^T \xi_0)^2}{\sigma^2} - \frac{(y^T \xi_0)^2}{2\sigma^2}$$

$$\Rightarrow \boxed{\log \Lambda_G(y) = \frac{1}{2\sigma^2} (y^T \xi_0)^2}$$

On a donc : $\forall y \in \mathbb{R}^N$, $\varphi_{GLRT}(y) = \begin{cases} 1 \text{ si } |y^T \xi_0| > \lambda \\ 0 \text{ sinon} \end{cases}$

où λ est un seuil à déterminer de manière à ce que $\alpha_{GLRT} = \beta_{\varphi_{GLRT}}(0) = \alpha$. On a alors :

$\alpha_{GLRT} = \mathbb{P}\left[|Y_0^T \xi_0| > \lambda \right]$ pour $Y_0 \sim \mathcal{N}(0, \sigma^2 I_N)$. On a donc

$$\alpha_{GLRT} = 1 - \mathbb{P}\left[|Y_0^T \xi_0| \leq \lambda \right]$$
$$= 1 - \mathbb{P}\left[-\lambda \leq Y_0^T \xi_0 \leq \lambda \right]$$

$Y_0^T \Theta_0$ est gaussien (transformation linéaire d'un vecteur gaussien).

$\mathbb{E}\left[Y_0^T \xi_0 \right] = 0$. $\mathrm{Var}\left\{ Y_0^T \xi_0 \right\} = \mathbb{E}\left[\xi_0^T Y_0 Y_0^T \xi_0 \right]$

$$= \xi_0^T \, \mathbb{E}\left[Y_0 Y_0^T \right] \xi_0$$
$$= \xi_0^T \sigma^2 I_N \xi_0 = \sigma^2 \|\xi_0\|_2^2$$

On a donc $\boxed{Y_0^T \xi_0 \sim \mathcal{N}(0, \sigma^2)}$ puisque $\|\xi_0\| = 1$.

$$\alpha_{GLRT} = 1 - \mathbb{P}\left[-\frac{\lambda}{\sigma} \leq \frac{Y_0^T \xi_0}{\sigma} \leq \frac{\lambda}{\sigma} \right]$$
$$= 1 - \left(\Phi\left(\frac{\lambda}{\sigma}\right) - \Phi\left(-\frac{\lambda}{\sigma}\right) \right)$$

$$Q\left(\frac{\lambda}{\sigma}\right) = 1 - \Phi\left(\frac{\lambda}{\sigma}\right) \quad \text{et} \quad \Phi\left(-\frac{\lambda}{\sigma}\right) = 1 - \Phi\left(\frac{\lambda}{\sigma}\right) = Q\left(\frac{\lambda}{\sigma}\right)$$

où Φ est la fonction de répartition de la loi $\mathcal{N}(0,1)$.

Donc $\boxed{\alpha_{GLRT} = 2Q\left(\frac{\lambda}{\sigma}\right)}$

On choisira $\boxed{\lambda = \sigma Q^{-1}(\gamma/2)}$. Finalement, on a :

$$\boxed{\forall y \in \mathbb{R}^N, \ \varphi_{GLRT}(y) = \begin{cases} 1 & \text{si } |y^T \xi_0| > \sigma Q^{-1}(\gamma/2) \\ 0 & \text{sinon} \end{cases}}$$

Interprétation de ce test : on projette l'observée sur la droite de vecteur directeur ξ_0. Et comme le signe de A n'est pas connu, on teste si l'amplitude de cette projection est suffisament grande. La projection sur la direction ξ_0 équivaut à une corrélation.

On calcule maintenant la puissance $\beta_{GLRT} = \beta_{\varphi_{GLRT}}(1)$. On a :

$$\beta_{GLRT} = \mathbb{P}\left[|Y_1^T \xi_0| > \lambda\right] \quad \text{où } Y_1 \sim \mathcal{N}(A\xi_0, \sigma^2 I_N).$$

$$= 1 - \mathbb{P}\left[-\lambda \leq Y_1^T \xi_0 \leq \lambda\right]$$

$Y_1^T \xi_0$ est gaussien en tant que transformation linéaire de Y_1.

$$\mathbb{E}\left[Y_1^T \xi_0\right] = \mathbb{E}\left[Y_1^T\right]\xi_0 = A\xi_0^T\xi_0 = A\|\xi_0\|^2 = A.$$

$\mathrm{var}\{Y_1^T \xi_0\} = \mathrm{var}\{(Y_1 - \mathbb{E}[Y_1])^T \xi_0\} = \mathrm{var}\{Z^T \xi_0\}$

avec $Z = Y_1 - \mathbb{E}[Y_1] \sim \mathcal{N}(0, \sigma^2 I_N)$. Donc :

$\mathrm{var}\{Y_1^T \xi_0\}$. Finalement, $Y_1^T \xi_0 \sim \mathcal{N}(A, \sigma^2)$. Aussi :

$$\beta_{GLRT} = 1 - \mathbb{P}\left[\frac{-\lambda - A}{\sigma} \leq \frac{Y_1^T \xi_0 - A}{\sigma} \leq \frac{\lambda - A}{\sigma}\right]$$

$$= 1 - \left(\Phi\left(\frac{\lambda - A}{\sigma}\right) - \Phi\left(-\frac{\lambda + A}{\sigma}\right)\right)$$

$$= Q\left(\frac{\lambda - A}{\sigma}\right) + Q\left(\frac{\lambda + A}{\sigma}\right)$$

Comme $\frac{\lambda}{\sigma} = Q^{-1}(\gamma/2)$, on a :

$$\boxed{\beta_{GLRT} = Q\left(Q^{-1}\left(\frac{\gamma}{2}\right) - \frac{A}{\sigma}\right) + Q\left(Q^{-1}\left(\frac{\gamma}{2}\right) + \frac{A}{\sigma}\right)}$$

Le lemme de Neyman-Pearson implique que :

$$\beta_{GLRT} \leq \beta^*$$

En comparant β_{GLRT} à β^*, on constatera que la différence est de l'ordre de 0.5 dB (si je ne me trompe pas)

Les tests GLRT ne sont généralement pas optimaux au sens de critère d'optimalité. Ils sont souvent utilisés lorsqu'on n'a pas de solutions issues des critères classiques ou qu'on ne voit pas de critère pertinent pour un problème donné. Mais souvent, ces tests ont des performances plutôt bonnes, comme dans le TP. Dans certains cas, on peut trouver en quel sens le test GLRT est optimal.