

Communications Numériques I

Module MH4 - cours 4/5

Pierre COMON

Résumé du cours 3

■ Distributions de probabilité

$$p_{\mathbf{z}}(\mathbf{v}) = p_{\mathbf{y}}(\mathbf{u}) \cdot |\det \mathbf{J}|$$

■ Circularité à l'ordre 2

$$\mathbb{E}[\mathbf{z}] = \mathbf{0}, \quad \mathbb{E}[\mathbf{z} \mathbf{z}^T] = \mathbf{0}$$

■ Fonctions de répartition

$$x \sim \mathcal{N}(m, \sigma^2) \implies 1 - F_x(a) = Q\left(\frac{a - m}{\sigma}\right), \quad a \geq m$$

■ Bornes

- Union: $\text{Max}_i \text{Pr}(E_i) \leq \text{Pr}(\cup_i E_i) \leq \sum_i \text{Pr}(E_i)$
- Chernoff: $\text{Pr}(X \geq \delta) \leq e^{-v\delta} \mathbb{E}[e^{vX}]$, $\mathbb{E}[x] = 0$, $\forall v$

■ D.s.p. de signaux modulés linéairement

$$z(t) \text{ stationnarisé} \implies S_z(f) = G(f) S_x(fT) G(f)^H$$

■ Karhunen-Loève

$$y(t) = \sum_i A_i f_i(t), \quad f_i \perp f_j$$

■ Filtre adapté en bruit blanc

$$h(t) = \frac{2}{N_o} (c \star g)(-t)$$

d.s.p. pour modulations linéaires

La densité spectrale de puissance dépend essentiellement du filtre d'émission/réception. Si x_k décorrelés:

$$S_z(f) = \frac{E_x}{T} |G(f)|^2$$

- Pour QAM, PSK, $\pi/4$ -DQPSK, etc... Par exemple $g(t)$ est un cosinus sur-élevé.
- Pour OQPSK, QAM, PSK, etc... $g(t)$ peut être une porte, et

$$|G(f)|^2 = \sqrt{T} \frac{\sin \pi f T}{\pi f T}$$

- Pour MSK, $\cos(\pi t/T)\Pi(t)$ et $\sin(\pi t/T)\Pi(t - T/2)$ ont même

$$|G(f)|^2 = \frac{\sqrt{8T}}{\pi} \frac{\cos \pi f T}{1 - 4 f^2 T^2}$$

Optimisation du filtre d'émission

■ Hypothèses simplificatrices

- le filtre de réception $H(f)$ est le filtre adapté
- la réponse globale $P(f)$ est imposée: $H(f) C(f) G(f) = P(f)$
- $\|G(f)\|^2 = 1$

■ Alors on peut chercher $|G|$ pour toute f

•

$$H \propto \frac{C^* G^*}{S_w} \Rightarrow P = \frac{|C|^2 |G|^2}{S_w} \Rightarrow$$

• D'où

$$|G| \propto \frac{\sqrt{P S_w}}{|C|} \quad \text{et} \quad |H| \propto \frac{\sqrt{P}}{\sqrt{S_w}}$$

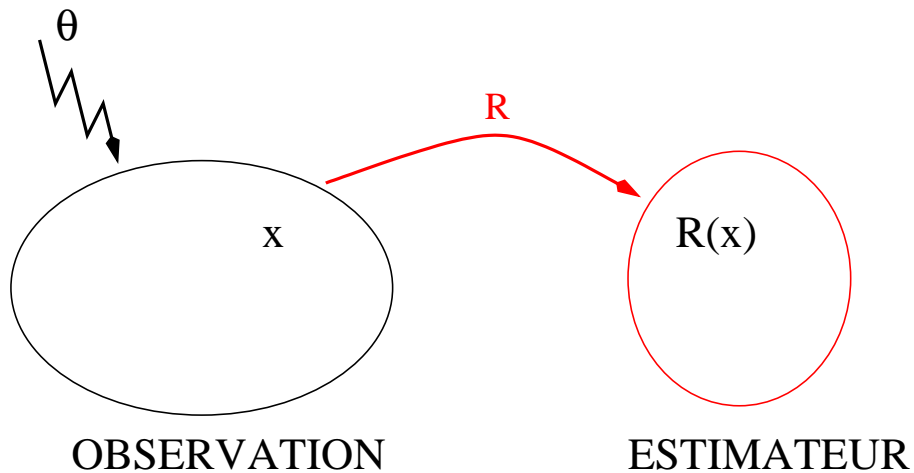
■ Si $C(f)$ et S_w sont constants, alors

$$|G(f)| \propto |H(f)| \propto \sqrt{P(f)}$$

■ En général, on prend pour $P(f)$ un cosinus sur-élevé $\Rightarrow G(f)$ est une “**racine de cosinus sur-élevé**”, à une phase près.

Théorie de la décision: rappels (1/4)

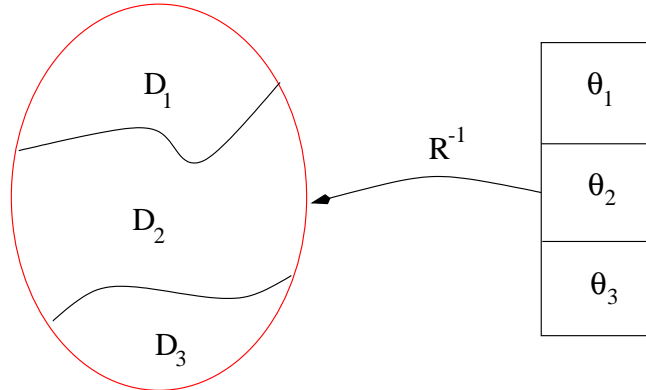
On cherche $\theta \in \Theta$



- R : application déterministe, $R(x)$ = estimateur de θ
- Détection/Classification: Θ = ensemble fini
- Estimation: Θ = ensemble éventuellement infini
- Région de confiance: Domaine $\mathcal{C}(x)$ où θ a de grandes chances de se trouver.

Théorie de la décision: rappels (2/4)

Domaines de décision



- q Hypothèses, $1 \leq i \leq q$
 - H_i : “La vraie valeur de θ est θ_i ”
 - $P_i = Pr(H_i)$
- Matrice de confusion: $C_{ij} = \int_{D_j} p_x(x|H_i) dx$
- Proba d’erreur totale: $P_\varepsilon = \sum_i \sum_{j \neq i} P_i C_{ij}$
- Terminologie de la détection binaire
 - $Pr(x \in D_1|H_0)$: erreur de 1ère espèce, **niveau** du test, P_{FA}
 - $Pr(x \in D_0|H_1)$: erreur de 2ème espèce, P_{ND}
 - $\beta = Pr(x \in D_1|H_1) = 1 - Pr(x \in D_0|H_1)$:
Probabilité de détection P_D , **puissance** du test

Théorie de la décision: rappels (3/4)

Approche bayésienne

- Minimisation d'un risque

$$\mathcal{R} = \sum_{i=1}^q P_i \sum_{j=1}^q \kappa_{ij} C_{ji}$$

κ_{ij} : fonction coût. On prend en général $0 = \kappa_{ii} < \kappa_{ij}$, $i \neq j$

- Solution en \mathbf{x} :

$D_j =$ domaine où $\Phi_j(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{i \neq j} P_i \kappa_{ij} p_x(\mathbf{x}|H_i)$ est le plus petit de tous les $\Phi_k(\mathbf{x})$, $1 \leq k \leq q$

- Cas équipénalisé (très courant): $\kappa_{ij} = 1 - \delta_{ij}$
On Maximise la proba A Posteriori (**MAP**)

$$\phi_j(\mathbf{x}) = P_j p_x(\mathbf{x}|H_j)$$

- Cas équadistribué: $P_j = 1/q$, $\forall q$
On maximise la proba a priori (**MV**)

$$\phi_j(\mathbf{x}) = p_x(\mathbf{x}|H_j)$$

- Cas binaire

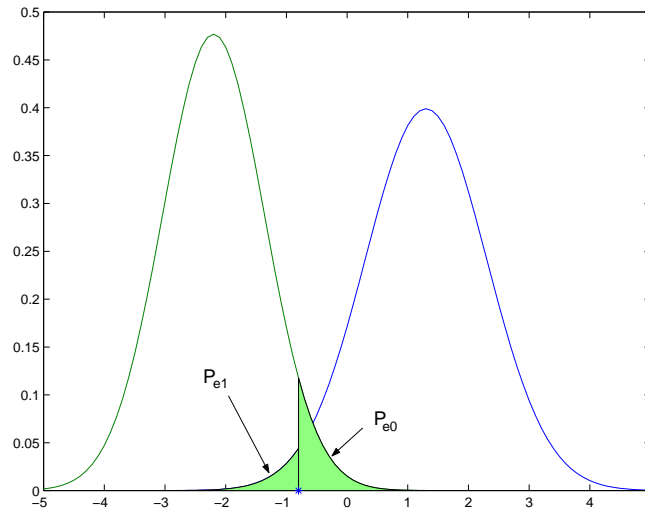
$$\frac{p_x(\mathbf{x}|H_1)}{p_x(\mathbf{x}|H_2)} \stackrel{D_1}{>} \frac{P_2 \kappa_{21}}{P_1 \kappa_{12}}$$

Théorie de la décision: rappels (4/4)

Cas particuliers importants

■ Cas binaire unimodal, détection sur la moyenne, $i \in \{0, 1\}$

- H_0 : $x < \eta$
- H_1 : $x > \eta$



■ Probabilités d'erreur

$$P_{\varepsilon 1} = Pr(x < \eta | H_1) = F_{x|H_1}(\eta)$$

$$P_{\varepsilon 0} = Pr(x > \eta | H_0) = 1 - F_{x|H_0}(\eta)$$

■ Cas gaussien: $H_i : x \sim \mathcal{N}(m_i, \sigma_i^2)$

$$P_{\varepsilon 1} = 1 - Q\left(\frac{\eta - m_1}{\sigma_1}\right)$$

$$P_{\varepsilon 0} = Q\left(\frac{\eta - m_0}{\sigma_0}\right)$$

Détection binaire gaussienne générale

$$H_i : \mathbf{x} \sim \mathcal{N}(\mathbf{m}_i, \mathbf{\Gamma}_i)$$

- Rapport de vraisemblance

$$\Lambda(\mathbf{x}) = \frac{\det \mathbf{\Gamma}_1^{1/2}}{\det \mathbf{\Gamma}_2^{1/2}} \cdot \frac{\exp[-(\mathbf{x} - \mathbf{m}_2)^H \mathbf{\Gamma}_2^{-1} (\mathbf{x} - \mathbf{m}_2)/2]}{\exp[-(\mathbf{x} - \mathbf{m}_1)^H \mathbf{\Gamma}_1^{-1} (\mathbf{x} - \mathbf{m}_1)/2]} \stackrel{D_2}{>} \Lambda_o$$

- log monotone \Rightarrow on peut prendre la statistique

$$(\mathbf{x} - \mathbf{m}_2)^H \mathbf{\Gamma}_2^{-1} (\mathbf{x} - \mathbf{m}_2) - (\mathbf{x} - \mathbf{m}_1)^H \mathbf{\Gamma}_1^{-1} (\mathbf{x} - \mathbf{m}_1) \stackrel{D_2}{>} \eta$$

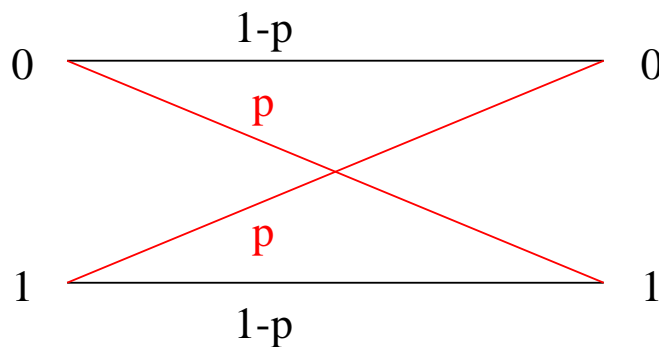
- Si $\mathbf{\Gamma}_1 = \mathbf{\Gamma}_2$, alors on pose $\mathbf{d} = \mathbf{m}_2 - \mathbf{m}_1$ et le **récepteur est linéaire** en \mathbf{x} :

$$\mathbf{d}^H \mathbf{\Gamma}^{-1} \mathbf{x} \stackrel{D_2}{>} \eta'$$

- Si $\mathbf{m}_1 = \mathbf{m}_2$, alors, le récepteur reste quadratique mais se simplifie:

$$\mathbf{x}^H [\mathbf{\Gamma}_1^{-1} - \mathbf{\Gamma}_2^{-1}] \mathbf{x} \stackrel{D_2}{>} \eta''$$

Canal binaire symétrique



- bit émis: \mathbf{u} , $H_i = Pr(\mathbf{u} = i)$; bit reçu: \mathbf{r}
- décision prise: $d(\mathbf{r})$. Décision bayésienne:

$$d(\mathbf{r}) = Arg \operatorname{Max}_{u \in \{0,1\}} P_u Pr(\mathbf{r} | H_u)$$

- $d(0) = 0 \Leftrightarrow P_0(1-p) > P_1 p$
 - $d(0) = 1 \Leftrightarrow P_0(1-p) < P_1 p$
 - $d(1) = 0 \Leftrightarrow P_1(1-p) < P_0 p$
 - $d(1) = 1 \Leftrightarrow P_1(1-p) > P_0 p$
- Cas particuliers
- $p \approx 0 \Rightarrow$ Proba d'erreur = $p(P_0 + P_1) = p \approx 0$
 - $p \approx 1 \Rightarrow$ Proba d'erreur = $(1-p)(P_0 + P_1) = 1-p \approx 0$,
mais le canal inverse les bits
 - $P_0 < p < 1/2 \Rightarrow d(\mathbf{u}) = 1, \forall \mathbf{u} \Rightarrow$ Proba d'erreur = P_0
 P_1 est trop forte, et seul l'*a priori* compte.

Théorie de l'estimation: rappels (1/2)

Approche bayésienne

- Minimisation du risque

$$\mathcal{R} = \mathbb{E}\{C[R(\mathbf{x}) - \boldsymbol{\theta}]\} = \int p_{x,\theta}(\mathbf{u}, \mathbf{v}) \int C[R(\mathbf{u}) - \mathbf{v}] d\mathbf{u} d\mathbf{v}$$

- $C[\cdot]$ est la fonction coût
 - $R(\mathbf{x})$ est l'estimateur de la variable $\boldsymbol{\theta}$
- Solution générale:

$$R(\mathbf{x}) = \mathit{Arg} \mathit{Min}_{\hat{\boldsymbol{\theta}}} \int C[\hat{\boldsymbol{\theta}} - \mathbf{v}] p_{\theta|x}(\mathbf{v}) d\mathbf{v}$$

Théorie de l'estimation: rappels (2/2)

- Si θ est uniformément distribuée
⇒ Solution du maximum de vraisemblance (**MV**)

$$R(\mathbf{x}) = \mathit{Arg} \mathit{Max}_{\mathbf{u}} p_{x|\theta}(\mathbf{u})$$

- Si le coût est uniforme: $C[\mathbf{u}] = 1 - \prod_i \delta(u_i)$
⇒ Solution du Maximum A Posteriori (**MAP**)

$$R(\mathbf{x}) = \mathit{Arg} \mathit{Max}_{\mathbf{v}} p_{\theta|x}(\mathbf{v})$$

- Si le coût est quadratique: $C[\mathbf{u}] = \|\mathbf{u}\|_2^2$ (norme L^2)
⇒ la solution est la moyenne a posteriori:

$$R(\mathbf{x}) = \mathit{E}\{\boldsymbol{\theta}|\mathbf{x}\}$$

- Si le coût est linéaire: $C[\mathbf{u}] = \|\mathbf{u}\|_1$ (norme L^1)
⇒ la solution est la médiane a posteriori

$$R(\mathbf{x}) = \mathit{mediane}\{\boldsymbol{\theta}|\mathbf{x}\}$$

Détection cohérente (1/2)

Formalisation générale

- Observation: $\mathbf{r} = \mathbf{A}\mathbf{x} + \mathbf{e}$; $\mathbf{e} \sim \mathcal{N}(0, \mathbf{\Gamma}_e)$,
 \mathbf{A} matrice connue

- En détection, \mathbf{x} est discrète. On pose donc $P_u = Pr(\mathbf{x} = \mathbf{u})$.

$$\mathbf{x}_{MAP} = \underset{\mathbf{u}}{Arg \text{ Max}} P_u \exp \left[-(\mathbf{r} - \mathbf{A}\mathbf{u})^H \mathbf{\Gamma}_e^{-1} (\mathbf{r} - \mathbf{A}\mathbf{u}) / 2 \right]$$

avec $\mathbf{u} \in$ alphabet fini. Une autre écriture utile est:

$$\begin{aligned} \mathbf{x}_{MAP} &= \underset{\mathbf{u}}{Arg \text{ Min}} \left\{ (\mathbf{r} - \mathbf{A}\mathbf{u})^H \mathbf{\Gamma}_e^{-1} (\mathbf{r} - \mathbf{A}\mathbf{u}) - 2 \log P_u \right\} \\ &\stackrel{\text{def}}{=} \underset{\mathbf{u}}{Arg \text{ Max}} \left\{ \langle \mathbf{r}, \mathbf{A}\mathbf{u} \rangle_{\mathbf{\Gamma}_b} - \|\mathbf{A}\mathbf{u}\|_{\mathbf{\Gamma}_b}^2 + 2 \log P_u \right\} \end{aligned}$$

- **NB:** En estimation, on aurait

- $\mathbf{x}_{MAP} = [\mathbf{A}^H \mathbf{\Gamma}_e^{-1} \mathbf{A} + \mathbf{\Gamma}_x^{-1}]^{-1} \mathbf{A}^H \mathbf{\Gamma}_e^{-1} \mathbf{r}$ si $\mathbf{x} \sim \mathcal{N}(0, \mathbf{\Gamma}_x)$
On montre aussi à l'aide d'un lemme d'inversion que

$$\mathbf{x}_{MAP} = \mathbf{\Gamma}_x \mathbf{A}^H \mathbf{\Gamma}_r^{-1} \mathbf{r}, \quad \mathbf{\Gamma}_r = \mathbf{\Gamma}_e + \mathbf{A} \mathbf{\Gamma}_x \mathbf{A}^H$$

- $\mathbf{x}_{MV} = [\mathbf{A}^H \mathbf{\Gamma}_e^{-1} \mathbf{A}]^{-1} \mathbf{A}^H \mathbf{\Gamma}_e^{-1} \mathbf{r}$

Détection cohérente (2/2)

- **Cas d'un signal x scalaire.** Alors

$$\mathbf{r} = \mathbf{a} x + \mathbf{e}$$

où $\mathbf{a} = \mathbf{vec}\{a_i\}$ est alors un vecteur, et u est scalaire.

La décision prise se simplifie:

$$x_{MAP} = \underset{u}{Arg \text{Max}} \left\{ 2u \langle \mathbf{r}, \mathbf{a} \rangle_{\Gamma_b} - |u|^2 \|\mathbf{a}\|_{\Gamma_b}^2 + 2 \log P_u \right\}$$

- **Cas d'une matrice Γ_b diagonale, et d'un signal x scalaire**

Alors $\mathbf{\Gamma}_b = \text{diag}\{\sigma_i^2\}$ et:

$$x_{MAP} = \underset{u}{Arg \text{Max}} \left\{ u \sum_k \frac{r_k^* a_k}{\sigma_k^2} - |u|^2 \sum_k \frac{|a_k|^2}{\sigma_k^2} + 2 \log P_u \right\}$$

→ quelquefois appelé “Maximum Ratio Combiner”