TD 4 : Calcul des estimateurs

Exercice 1 : Estimation par moindres carrés

1. Soient les vecteurs colonne $\underline{\boldsymbol{\theta}}$ et $\underline{\boldsymbol{b}}$ de N lignes et la matrice $\mathbf{A}$ de taille $N \times N$, symétrique ($\mathbf{A}^T = \mathbf{A}$), définie positive ($\underline{\boldsymbol{u}}^T \mathbf{A} \underline{\boldsymbol{u}} \geq 0, \forall \underline{\boldsymbol{u}}$), inversible (avec $\mathbf{A}^{-1}$ tel que $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$). Montrer que :
$$\underline{\boldsymbol{\theta}}^T \mathbf{A} \underline{\boldsymbol{\theta}} - 2\underline{\boldsymbol{\theta}}^T \underline{\boldsymbol{b}} = (\underline{\boldsymbol{\theta}} - \underline{\boldsymbol{\theta}}_0)^T \mathbf{A} (\underline{\boldsymbol{\theta}} - \underline{\boldsymbol{\theta}}_0) - \underline{\boldsymbol{\theta}}_0^T \mathbf{A} \underline{\boldsymbol{\theta}}_0 \quad \text{avec } \underline{\boldsymbol{\theta}}_0 = \mathbf{A}^{-1} \underline{\boldsymbol{b}}$$
En déduire que le minimiseur du critère $J(\underline{\boldsymbol{\theta}}) = \underline{\boldsymbol{\theta}}^T \mathbf{A} \underline{\boldsymbol{\theta}} - 2\underline{\boldsymbol{\theta}}^T \underline{\boldsymbol{b}} + c$ est $\boldsymbol{\theta}_0 = \mathbf{A}^{-1}\underline{\boldsymbol{b}}$.
2. Soit le critère $J(\underline{\boldsymbol{\theta}}) = \left(\underline{\boldsymbol{y}} - \mathbf{R}\underline{\boldsymbol{\theta}}\right)^T \mathbf{Q} \left(\underline{\boldsymbol{y}} - \mathbf{R}\underline{\boldsymbol{\theta}}\right)$ à minimiser par rapport au paramètre $\underline{\boldsymbol{\theta}}$, avec $\mathbf{Q}$ une matrice symétrique définie positive. Montrer en exploitant la réponse à la question précédente que le minimum est atteint pour $\underline{\boldsymbol{\theta}}_0 = \left(\mathbf{R}^T \mathbf{Q} \mathbf{R}\right)^{-1} \mathbf{R}^T \mathbf{Q} \underline{\boldsymbol{y}}$
3. Régression polynomiale : on cherche à estimer les paramètres $\underline{\boldsymbol{\theta}} = [\alpha, \beta, \gamma, \delta]^T$ tel que le polynôme $P(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta$ passe « au mieux » dans le nuage de points $\{x_n, y_n\}_{n=1\ldots N}$ au sens du critère quadratique. Écrire le critère à minimiser et déduire des réponses aux questions précédentes l'expression de l'estimateur $\overset{o}{\underline{\boldsymbol{\theta}}}$.
4. Dans le cas précédent, si l'on suppose que les données ont été formées par le modèle $y_n = \alpha x_n^3 + \beta x_n^2 + \gamma x_n + \delta + \epsilon_n$ avec les perturbations ϵ_n considérées comme aléatoires indépendantes identiquement distribuées suivant une loi gaussienne $\mathcal{N}(0, \sigma_\epsilon^2)$. Donner l'expression de l'estimateur du maximum de vraisemblance de $\underline{\boldsymbol{\theta}}$.
5. Toujours dans le cas précédent, on suppose maintenant que les perturbations $\boldsymbol{\epsilon}$ ne sont plus indépendantes mais toujours distribuées suivant une loi gaussienne $\mathcal{N}(\underline{\mathbf{0}}, \mathbf{\Gamma}_\epsilon)$. Donner l'expression de l'estimateur du maximum de vraisemblance de $\underline{\boldsymbol{\theta}}$.
6. Calculer le biais et la matrice de covariance de l'estimateur de la question précédente.

Exercice 2 : Biais et variance de l'estimateur des Moindres carrés (TP 1)

On se place dans le cas d'un modèle LP et de perturbations additives gaussiennes centrées de matrice de covariance $\mathbf{\Gamma}_E$:
$$\underline{\boldsymbol{y}}(\underline{\boldsymbol{\theta}}) = R\underline{\boldsymbol{\theta}} + \underline{\boldsymbol{\epsilon}}$$

1. Montrer que l'estimateur du maximum de vraisemblance est à biais nul. et de matrice de covariance :
$$V_{\overset{o}{\underline{\boldsymbol{\theta}}}_{\mathrm{MV}}}(\underline{\boldsymbol{\theta}}) \triangleq \mathrm{E}_{Y|\theta}\{(\overset{o}{\underline{\boldsymbol{\theta}}}_{\mathrm{MV}} - \underline{\boldsymbol{\theta}})(\overset{o}{\underline{\boldsymbol{\theta}}}_{\mathrm{MV}} - \underline{\boldsymbol{\theta}})^T | \boldsymbol{\theta}\} = \left(R^T \mathbf{\Gamma}_E^{-1} R\right)^{-1}$$
qui vérifie donc l'égalité de la borne de Cramer-Rao.

Exercice 3 : Modèle convolutif, perturb. additives (TP 1)

De nombreux systèmes peuvent être modélisés par une relation entrée/sortie de type convolution :
$$y_{\mathcal{M}}[n] = \sum_{k=0}^{M-1} h[k]x[n-k],$$
où h est la réponse impulsionnelle du système de durée finie M. C'est par exemple le cas en télé-communication où h modélise le canal de transmission et en géophysique pour la caractérisation de la structure du sol par échographie.
Trois types de problèmes peuvent alors se poser :

— Lorsqu'on maîtrise le signal d'entrée et que l'on désire *identifier* le système en estimant sa réponse impulsionnelle $h[n]$;
— lorsqu'on connaît la réponse impulsionnelle $h[n]$ et que l'on souhaite *déconvoluer* le signal $x[n]$;
— lorsqu'on ne connaît ni la réponse impulsionnelle $h[n]$ ni l'entrée $x[n]$, on parle alors de *déconvolution aveugle.*

On va ici s'intéresser aux deux premières situations, la dernière étant beaucoup plus délicate. On considère le cas ou les perturbations sont additives gaussiennes *i.i.d.* centrées de variance connue σ^2.

1. Dans le cas de l'identification, on peut écrire la relation entre les paramètres à estimer $\underline{\boldsymbol{h}} = [h[0], h[1], \ldots h[M-1]]^T$ et les données observées $\underline{\boldsymbol{y}} = [y[0], y[1], \ldots y[N-1]]^T$ sous forme matricielle : $\underline{\boldsymbol{y}} = X\underline{\boldsymbol{h}} + \underline{\boldsymbol{\epsilon}}$. Ecrire la matrice X, quelle est sa dimension ?
2. Écrire le critère du maximum de vraisemblance pour $\underline{\boldsymbol{h}}$ ainsi que la valeur de $\overset{o}{\underline{\boldsymbol{h}}}$ minimisant ce critère.
3. Dans le cas de la déconvolution, on peut également écrire la relation entre les paramètres à estimer $\underline{\boldsymbol{x}} = [x[0], x[1], \ldots x[N-1]]^T$ et les données observées $\underline{\boldsymbol{y}}$ sous forme matricielle : $\underline{\boldsymbol{y}} = H\underline{\boldsymbol{x}} + \underline{\boldsymbol{\epsilon}}$. Écrire la matrice H, quelle est sa dimension ? Pourquoi un critère du maximum de vraisemblance est-il mal adapté à l'estimation de $\underline{\boldsymbol{x}}$?
4. On considère un critère du maximum *a posteriori* dans le cas d'un modèle gaussien centré de matrice de covariance $\mathbf{\Gamma}_x$ pour $\underline{\boldsymbol{x}}$. Écrire le critère ainsi que la valeur de $\overset{o}{\underline{\boldsymbol{x}}}$ le minimisant.
5. On considère toujours un critère du maximum *a posteriori* dans le cas d'un champ de Gibbs gaussien de densité de probabilité $f_X(\underline{\boldsymbol{x}}) \propto \exp -\frac{1}{T} \sum_{i \sim j} (x[i] - x[j])^2$ où $i \sim j$ signifie i et j voisins c'est-à-dire de que la somme opère sur tous les couples de la forme $(x[i], x[i+1])$. Écrire la densité f_X sous la forme $f_X(\underline{\boldsymbol{x}}) \propto \exp -\frac{1}{T}(D\underline{\boldsymbol{x}})^T(D\underline{\boldsymbol{x}})$ avec une matrice D à préciser. Écrire le critère correspondant à ce modèle et en donner une interprétation physique. Donner la valeur de $\overset{o}{\underline{\boldsymbol{x}}}$ minimisant ce critère.

Exercice 4 : Modèle de bruit coloré et blanchiment (TP 2)

Soit un modèle de signal $y[n]$ constitué du filtrage d'un signal $x[n]$ par un filtre causal H auquel s'ajoute un bruit blanc (variables aléatoires indépendantes centrées et identiquement distribuées) $\epsilon[n]$: $y[n] = y_{\mathcal{M}}[n] + \epsilon[n]$, avec $y_{\mathcal{M}}[n] = \sum_{k=0}^{\infty} h[k]x[n-k]$.
On s'intéresse au cas où la réponse impulsionnelle du filtre vérifie l'équation de récurrence :

$$h[n] = a_1 h[n-1] + a_2 h[n-2] \text{ pour } n \geq 2,$$

les coefficients $\{a_1, a_2, h[0], h[1]\}$ étant connus.

1. Montrer que le signal non bruité $y_{\mathcal{M}}[n]$ satisfait une équation de récurrence d'ordre 2
$$y_{\mathcal{M}}[n] = a_1 y_{\mathcal{M}}[n-1] + a_2 y_{\mathcal{M}}[n-2] + b_0 x[n] + b_1 x[n-1] + b_2 x[n-2],$$
de coefficients $\{a_1, a_2, b_0, b_1, b_2\}$ à préciser, soit un filtrage par un filtre d'ordre 2, pour le signal d'entrée $x[n]$.
2. Montrer que le signal bruité $y[n]$ satisfait la même équation de récurrence que $y_{\mathcal{M}}[n]$ à laquelle s'ajoute un terme de bruit $b[n]$ correspondant au filtrage du bruit blanc $\epsilon[n]$ par un filtre G dont on précisera les coefficients. Ce type de bruit, non blanc, est souvent appelé bruit coloré.
3. Montrer qu'en construisant le signal $\widetilde{y}[n]$ par filtrage du signal $y[n]$ par le filtre inverse de G (de fonction de transfert $1/G(z)$), ce signal filtré vérifie la même équation de récurrence que le signal non bruité $y_{\mathcal{M}}[n]$ à laquelle s'ajoute le terme de bruit blanc $\epsilon[n]$. On dit qu'on a procédé à une étape de blanchiment du bruit.