Examen Master SICOM, Cours Optimisation
Luc Pronzato, janvier 2006

QUELQUES INDICATIONS

Les parties 1 et 2 sont totalement indépendantes.

On prendra soin de **souligner les vecteurs** et de noter les matrices par des lettres capitales. Un vecteur $\mathbf{v}$ est toujours supposé être du type colonne, on notera $\mathbf{v}^\top$ un vecteur ligne.

ÉNONCÉ DU PROBLÈME

PARTIE 1

On considère un modèle paramétrique dont la réponse $\eta(\theta,t)$ pour un vecteur de paramètres $\theta = (\theta_1,\theta_2)^\top$ en fonction du temps t est donnée par la solution de l'équation différentielle

$$\theta_1 \left[\frac{\partial^2 \eta(\theta,t)}{\partial t^2} + \eta(\theta,t) \right] + \frac{\partial \eta(\theta,t)}{\partial t} = \theta_2 u(t)\,, \tag{1}$$

avec des conditions initiales nulles. On souhaite estimer θ à partir d'observations $y(t_1),\ldots,y(t_n)$ recueillies aux instants $t_1,\ldots,t_n$ pour le critère des moindres carrés

$$J(\theta) = \frac{1}{2}\sum_{i=1}^{n}[y(t_i) - \eta(\theta,t_i)]^2\,.$$

1) Exprimez le gradient du critère $\mathbf{g}(\theta) = \partial J(\theta)/\partial\theta$ en fonction des fonctions de sensibilité

$$s_1(\theta,t) = \frac{\partial \eta(\theta,t)}{\partial \theta_1} \ \text{ et } \ s_2(\theta,t) = \frac{\partial \eta(\theta,t)}{\partial \theta_2}\,.$$

On notera dans la suite plus simplement $\eta = \eta(\theta,t)$, $s_1 = s_1(\theta,t)$, $s_2 = s_2(\theta,t)$, et $\dot\eta$, $\dot s_1$ $\dot s_2$, $\ddot\eta$, $\ddot s_1$ $\ddot s_2$ leurs dérivées premières et secondes par rapport à t.

2) Écrivez les équations différentielles satisfaites par s_1 et s_2.

3) Combien d'équations différentielles faut-il résoudre pour connaître η, s_1 et s_2 ? Écrivez ces équations et représentez les sur un schéma bloc (on notera $1/s$ le bloc intégrateur et on indiquera sur le schéma où sont obtenus η, s_1 et s_2).

4) En définissant $\mathbf{x}$ le vecteur d'état

$$\mathbf{x} = (\dot s_2, s_2, \dot s_1 s_1)^\top$$

écrivez les équations différentielles de la question 3 sous la forme

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u(t) \tag{2}$$

et explicitez la matrice $\mathbf{A} \in \mathbb{R}^4 \times \mathbb{R}^4$ et le vecteur $\mathbf{b} \in \mathbb{R}^4$. Exprimer $\mathbf{z} = (\eta, s_1, s_2)^\top$ sous la forme $\mathbf{z} = \mathbf{C}\mathbf{x}$ en explicitant la matrice $\mathbf{C}$.

5) Le calcul du déterminant $d = \det[\lambda\mathbf{I} - \mathbf{A}]$, avec $\mathbf{I}$ la matrice identité, donne

$$d = \frac{(\lambda^2\theta_1 + \lambda + \theta_1)^2}{\theta_1^2}\,.$$

Discutez la stabilité des équations différentielles de la question 3 suivant la valeur de θ_1.

6) Dans l'équation différentielle de départ (1), posez $\mathbf{x} = (\dot\eta, \eta)^\top$, et écrivez l'équation différentielle sous la forme (2) (où à présent $\mathbf{A} \in \mathbb{R}^2 \times \mathbb{R}^2$ et $\mathbf{b} \in \mathbb{R}^2$). Calculer $\det[\lambda\mathbf{I} - \mathbf{A}]$ et discuter la stabilité de (1) en fonction de θ_1. Pourquoi les résultats des calculs de la question 5) étaient-ils prévisibles ?

PARTIE 2

On considère l'algorithme du gradient à pas optimum pour l'optimisation de la fonction

$$f(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^\top \mathbf{A}(\mathbf{x} - \mathbf{x}^*)$$

avec $\mathbf{x} \in \mathbb{R}^d$ et $\mathbf{A}$ une matrice symétrique définie positive.

1) Calculez le gradient de f en $\mathbf{x}$, que l'on notera $\mathbf{g}(\mathbf{x})$.

2) Si $\mathbf{x}_k$ est la valeur proposée par l'algorithme à l'itération k, donnez la valeur optimale du pas, notée γ_k, pour passer au point suivant $\mathbf{x}_{k+1}$. Exprimez γ_k en fonction de $\mathbf{A}$ et $\mathbf{g}_k = \mathbf{g}(\mathbf{x}_k)$ puis $\mathbf{x}_{k+1}$ en fonction de $\mathbf{x}_k$, $\mathbf{A}$ et $\mathbf{g}_k$.

3) Donnez l'expression de $\mathbf{g}_{k+1}$ en fonction de $\mathbf{A}$ et $\mathbf{g}_k$.

4) On définit

$$\mathbf{z}_k = \frac{\mathbf{g}_k}{\sqrt{\mathbf{g}_k^\top \mathbf{g}_k}}$$

et $\mu_i^k = \mathbf{z}_k^\top \mathbf{A}^i \mathbf{z}_k$ pour tout i. Que vaut μ_0^k ? Exprimez $\mathbf{z}_{k+1}$ en fonction de $\mathbf{A}$, μ_1^k, μ_2^k et $\mathbf{z}_k$. Exprimez ensuite μ_i^{k+1} en fonction de μ_i^k, μ_{i+1}^k, μ_{i+2}^k, μ_1^k, et μ_2^k. Calculez la vitesse de convergence à l'itération k

$$r_k = \frac{f(\mathbf{x}_{k+1})}{f(\mathbf{x}_k)}$$

en fonction de μ_1^k et μ_{-1}^k.

5) On suppose que $\mathbf{A}$ est sous la forme diagonale, $\mathbf{A} = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$. Expliquez pourquoi $(\{\mathbf{z}_k\}_i)^2$, avec $\{\mathbf{z}_k\}_i$ la composante numéro i de $\mathbf{z}_k$, peut être considéré comme une masse affectée à la valeur propre λ_i de $\mathbf{A}$ à l'itération k. Quelle est alors l'interprétation des μ_i^k ? Montrez que

$$(\{\mathbf{z}_{k+1}\}_i)^2 = H_k(\lambda_i, \mu_1^k, \mu_2^k)(\{\mathbf{z}_k\}_i)^2$$

pour tout $i = 1, \ldots, d$, et explicitez la fonction H_k.