

Examen - Modélisation - Estimation Paramétrique

L'usage des calculatrices est interdit. Seule une feuille de notes manuscrites A4 est autorisée. Les parties sont indépendantes. Il est conseillé de lire attentivement le sujet avant de commencer. Les réponses devront être soigneusement argumentées.

1 Gradient, Newton, et Gauss-Newton

On considère la fonction coût suivante

$$c(\mathbf{p}) = ((p_1 - p_2) - 1)^2 + ((p_1 - 2p_2) - 0.5)^2. \quad (1)$$

On utilise une méthode de recherche itérative du minimum de $c(\mathbf{p})$. Pour cela, on part du point $\mathbf{p}_0 = (1, 1)^T$.

1. Méthode du gradient
 - (a) Donner la direction de descente \mathbf{d}_g donnée par cette méthode.
 - (b) Evaluer $c(\mathbf{p}_1)$ avec $\mathbf{p}_1 = \mathbf{p}_0 + \mathbf{d}_g$.
2. Méthode de Newton.
 - (a) Donner la direction de descente \mathbf{d}_n donnée par cette méthode.
 - (b) Evaluer $c(\mathbf{p}_1)$ avec $\mathbf{p}_1 = \mathbf{p}_0 + \mathbf{d}_n$.
3. Méthode de Gauss-Newton.
 - (a) Donner la direction de descente \mathbf{d}_{gn} donnée par cette méthode.
 - (b) Evaluer $c(\mathbf{p}_1)$ avec $\mathbf{p}_1 = \mathbf{p}_0 + \mathbf{d}_{gn}$.
4. Conclusion.

2 Estimation Min-Max

On considère le modèle à temps discret suivant

$$y_m(\boldsymbol{\theta}, k) = \theta_1 y(k-1) + \theta_2 y(k-2) \quad (2)$$

dont l'objectif est de représenter les sorties $y(k)$ mesurées aux instants $k = 1, \dots, N$. L'objectif est d'obtenir une estimée min-max

$$\hat{\boldsymbol{\theta}}_{\text{mM}} = \arg \min_{\boldsymbol{\theta}} \max_{k=3, \dots, N} |y(k) - y_m(\boldsymbol{\theta}, k)|$$

du vecteur de paramètres $\boldsymbol{\theta}$.

1. Le modèle (2) est-il linéaire ou non-linéaire en les paramètres ? Justifier.
2. Dans quelle situation est-il préférable d'utiliser une estimation de type min-max à la place d'une estimation au sens des moindres carrés ?
3. Montrer que $\hat{\boldsymbol{\theta}}_{\text{mM}}$ peut être obtenu à partir de la solution d'un programme linéaire de type

$$\min \mathbf{c}^T \mathbf{x}$$

$$\text{sous la contrainte } \mathbf{A}\mathbf{x} \leq \mathbf{b},$$

dont les éléments (\mathbf{b} , \mathbf{c} , \mathbf{x} , et \mathbf{A}) seront à spécifier.

3 Calcul de gradient par état adjoint

On considère le modèle sous forme de représentation d'état à temps discret suivant

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \end{pmatrix} = \begin{pmatrix} p_1 x_1(k) + p_2 x_2^2(k) \\ p_3 x_1^2(k) \end{pmatrix}$$
$$y_m(k) = x_1(k) + x_2(k)$$

avec

$$\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} p_4 \\ 0 \end{pmatrix}$$

d'un système dynamique dont on a mesuré la sortie $y(k)$ aux instants $k = 0, \dots, N$.

On souhaite réaliser l'estimée au sens des moindres carrés du vecteur de paramètres $\mathbf{p} = (p_1, \dots, p_4)^T$ à partir des mesures $y(k)$ aux instants $k = 0, \dots, N$.

1. Donner la forme du critère $c(\mathbf{p})$ à minimiser pour obtenir cette estimée.
2. Donner la dynamique à temps rétrograde satisfaite par l'état adjoint qui va permettre d'obtenir le gradient du critère $c(\mathbf{p})$. Préciser la valeur initiale de l'état adjoint.
3. Indiquer la manière dont est obtenu le gradient à partir de l'état adjoint.

4 Construction d'un modèle de comportement

On souhaite réaliser un système de téléopération d'un véhicule, d'un drone ou d'un train depuis un site distant, voir la Figure 1. Pour cela, afin de permettre à l'opérateur distant de réagir en cas d'urgence, le véhicule doit envoyer des informations visuelles sur son environnement avec une latence extrêmement faible.



FIGURE 1 – Opérateur télé-conduisant un véhicule distant (source <https://designateddriver.ai/technology/>)

Pour cela, le véhicule doit être équipé de dispositifs d'acquisition (caméra), de compression et de transmission de contenus vidéo efficaces. La latence de bout-en-bout est l'intervalle de temps entre l'instant auquel un événement apparaît en face de la caméra embarquée dans le véhicule et l'instant auquel cet événement est affiché sur l'écran de l'opérateur distant.

1. Quelle est la distance parcourue en 100 ms par un véhicule se déplaçant à 50 km/h ? En déduire un ordre de grandeur de la latence de bout-en-bout tolérable pour un système de télé-conduite.

La Figure 2 présente un schéma simplifié d'un dispositif d'acquisition, de compression et de transmission de contenus vidéo acquis depuis un véhicule. La latence de bout-en-bout est due

- au temps mis par la caméra pour acquérir l'image
- au temps mis par le codeur vidéo pour compresser l'image acquise
- au temps de séjour de l'image compressée dans le tampon d'émission, en attendant que la radio dispose d'opportunités de transmission des paquets dans lesquels l'image compressée a été stockée
- au temps de transmission des paquets par la radio et par le réseau de communication filaire qui suit l'antenne radio réceptrice
- au temps de stockage dans le tampon de réception des paquets reçus
- au temps mis par le décodeur vidéo pour décompresser les paquets reçus
- au temps d'affichage de l'image décompressée.

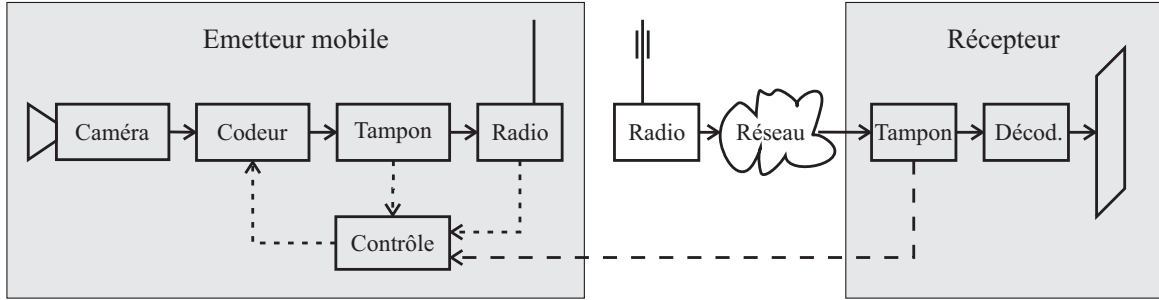


FIGURE 2 – Schéma d'une chaîne d'acquisition, compression et transmission de contenus vidéo acquis en temps réel

Comme le véhicule se déplace dans son environnement, la capacité¹ du canal de communication entre l'antenne radio émettrice sur le véhicule et l'antenne radio réceptrice (dans la station de base ou la gNodeB pour la 5G) varie fortement au cours du temps et en fonction du nombre d'utilisateurs connectés au réseau. Il est par conséquent indispensable d'ajuster régulièrement les paramètres du codeur vidéo afin que la quantité de bits générés par image (débit binaire) soit le plus proche possible de la capacité du canal de communication. C'est le rôle du bloc *contrôle* de la Figure 2.

Pour un dispositif de compression vidéo, le débit R_n en sortie de codeur mesure la quantité de bits à transmettre par seconde pour envoyer l'image n une fois qu'elle est compressée. Si T_i est l'intervalle de temps entre l'acquisition de deux images, le nombre de bits représentant l'image compressée n est $R_n T_i$. La distorsion D_n mesure la qualité de l'image n après compression et décompression. La distorsion correspond à la moyenne des carrés des différences entre l'intensité lumineuse des pixels de l'image acquise et de l'intensité lumineuse des pixels affichés au récepteur. Plus R_n est élevé, plus D_n est faible et meilleure est la qualité de l'image reçue.

2. Que se passe-t-il si le débit binaire R_n en sortie de codeur vidéo est trop important par rapport à la capacité ? Que se passe-t-il lorsqu'il est beaucoup plus faible que la capacité ?

Le compromis entre R_n et D_n dépend principalement du paramètre de quantification (QP) sélectionné par le codeur vidéo pour compresser l'image n . Ce paramètre sera noté Q_n dans ce qui suit. La relation entre R_n et D_n est complexe à décrire et fait l'objet de la suite de cet exercice où nous allons essayer d'en construire un modèle de comportement.

Pour la séquence vidéo ParkScene, la Figure 3 illustre l'évolution de R_n en fonction de $\log(D_{n-1})$ pour différentes valeurs de $Q_n \in \{20, 22, \dots, 42\}$. Les valeurs expérimentales correspondent aux points sur la Figure 3. Les 12 valeurs de Q_n sont notées $Q_n^{(i)}$, $i = 1, \dots, 12$. Pour chaque $Q_n^{(i)}$, 13 valeurs de distorsion de l'image précédente $n - 1$ ont été considérées. Elles sont notées $D_{n-1}^{(i,j)}$ et les valeurs obtenues de débit sont notées $R_n^{(i,j)}$, $i = 1, \dots, 12$ et $j = 1, \dots, 13$.

3. Est-ce que R_n dépend uniquement de Q_n ?

1. la quantité de bits qu'il est possible d'envoyer par seconde

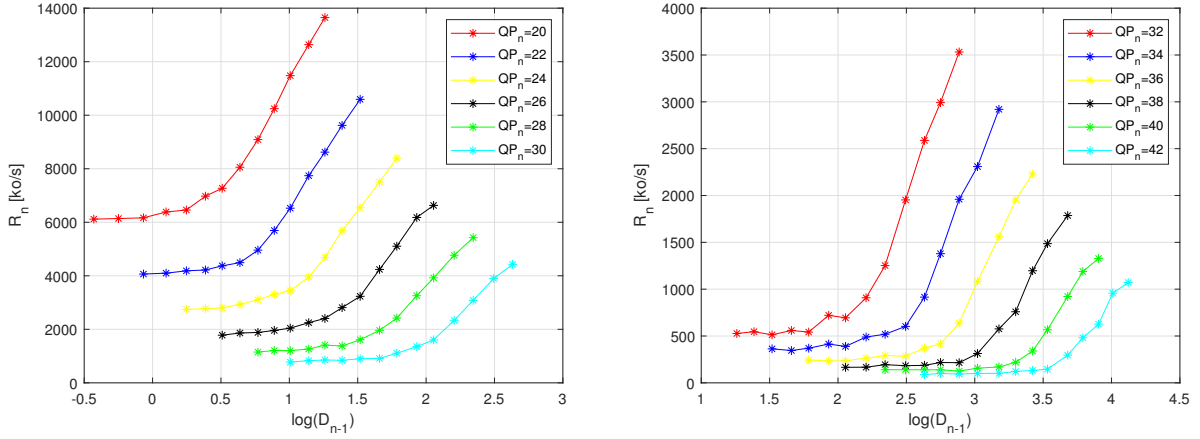


FIGURE 3 – Séquence vidéo *ParkScene* : Débit R_n en sortie de codeur vidéo pour l'image n en fonction du logarithme de la distorsion de l'image $n - 1$, pour différentes valeur du paramètre de quantification Q_n de l'image n

4. Que peut-on dire de la dépendance entre R_n et Q_n lorsque D_{n-1} est faible ? Que peut-on en dire lorsque D_{n-1} est élevé ?

Une première partie du modèle de comportement reliant R_n et Q_n considérera D_{n-1} très faible. La figure 4 illustre la valeur minimale $R_n^{(i,1)}$ de R_n obtenue pour chaque $Q_n^{(i)}$, $i = 1, \dots, 12$.

5. Lorsque D_{n-1} est très faible, en vous servant de la Figure 4, quelle structure de modèle

$$R_n = f_1(Q_n, \mathbf{p}_n)$$

pouvez-vous proposer ? Ce modèle est-il linéaire ou non-linéaire en les paramètres ?

6. Quelle est la dimension du vecteur \mathbf{p}_n ?
7. Quel type de critère peut-on considérer pour estimer \mathbf{p}_n à partir des $Q_n^{(i)}$, $i = 1, \dots, 12$ et des valeurs correspondantes de $R_n^{(i,1)}$? Donner l'expression de ce critère ?
8. Donner une valeur approximative de l'estimée $\hat{\mathbf{p}}_n$ de \mathbf{p}_n au sens du critère proposé.

La Figure 5 illustre les valeurs de $\underline{R}_n^{(i,j)} = R_n^{(i,j)} - f_1(Q_n^{(i)}, \hat{\mathbf{p}}_n)$, $i = 1, \dots, 12$ et $j = 1, \dots, 13$ pour l'image n en fonction du logarithme de la distorsion de l'image $n - 1$, pour différents $Q_n^{(i)}$, $i = 1, \dots, 12$.

9. On considère une valeur $Q_n^{(i)}$ du paramètre de quantification fixée. A partir de l'observation de la Figure 5, justifier le choix de la structure de modèle paramétrique suivante

$$\underline{R}_n = q_1 (1 + \tanh(q_2 (\log(D_{n-1}) - q_3))) \quad (3)$$

afin de relier \underline{R}_n à $\log(D_{n-1})$. Le vecteur des paramètres de ce modèle est $\mathbf{q} = (q_1, q_2, q_3)^T$.

10. Le modèle (3) est-il linéaire ou non-linéaire en les paramètres ?

Pour chaque valeur de $Q_n^{(i)}$, une estimation au sens des moindres carrés $\hat{\mathbf{q}}$ de \mathbf{q} est faite à partir de $\underline{R}_n^{(i,j)}$ et $D_{n-1}^{(j)}$, $j = 1, \dots, 13$. La Figure 6 illustre les estimées \hat{q}_1 , \hat{q}_2 et \hat{q}_3 obtenues pour différentes valeurs de $Q_n^{(i)}$ en fonction de Q_n (haut-gauche, bas-gauche, bas-droit) ou de $\log(Q_n)$ (haut-droit).

11. Peut-on dire que \mathbf{q} est indépendant de Q_n ?
12. A partir des courbes de la Figure 6, proposer une structure de modèle permettant de relier
 - (a) q_1 à Q_n ; elle sera notée $q_1 = g_1(Q_n, \mathbf{p})$ ou $q_1 = g_1(\log(Q_n), \mathbf{p})$;

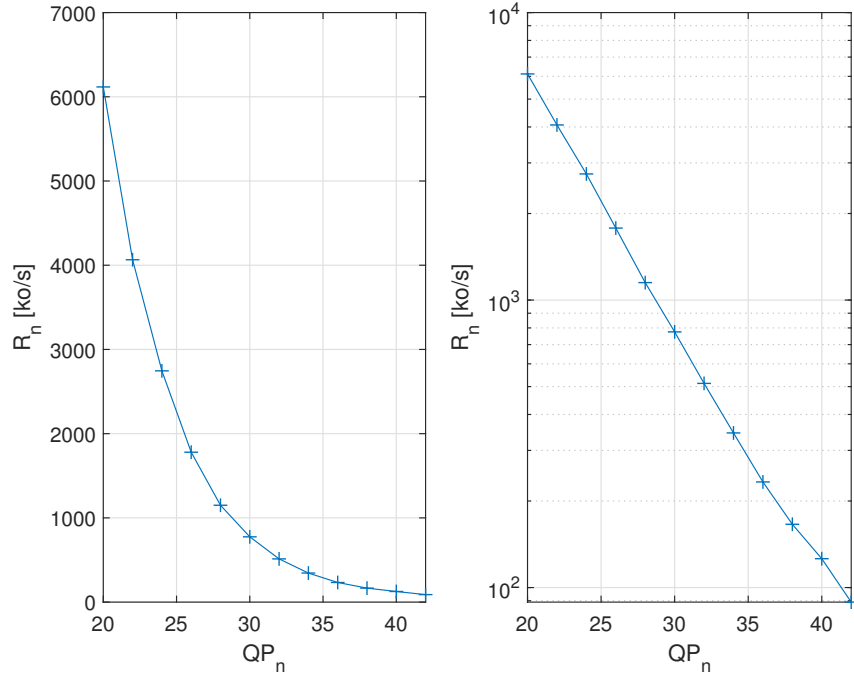


FIGURE 4 – Séquence vidéo *ParkScene* : Débit R_n minimal obtenu en sortie de codeur vidéo pour l'image n en fonction du paramètre de quantification Q_n de l'image n , échelle linéaire (à gauche) et logarithmique (à droite)

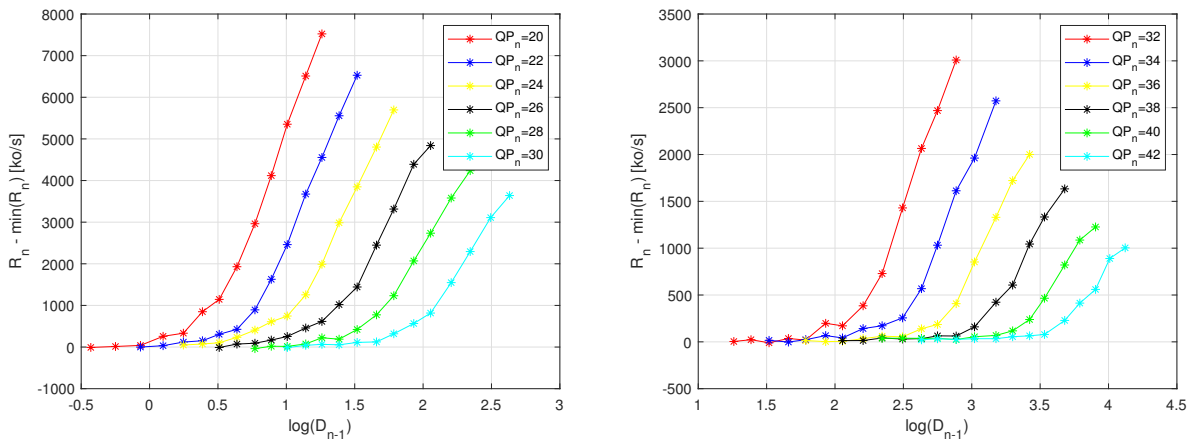


FIGURE 5 – Séquence vidéo *ParkScene* : valeurs de $R_n^{(i,j)} - f_1(Q_n^{(i)}, \hat{\mathbf{p}}_n)$ pour l'image n en fonction du logarithme de la distorsion de l'image $n - 1$, pour différents $Q_n^{(i)}$, $i = 1, \dots, 12$

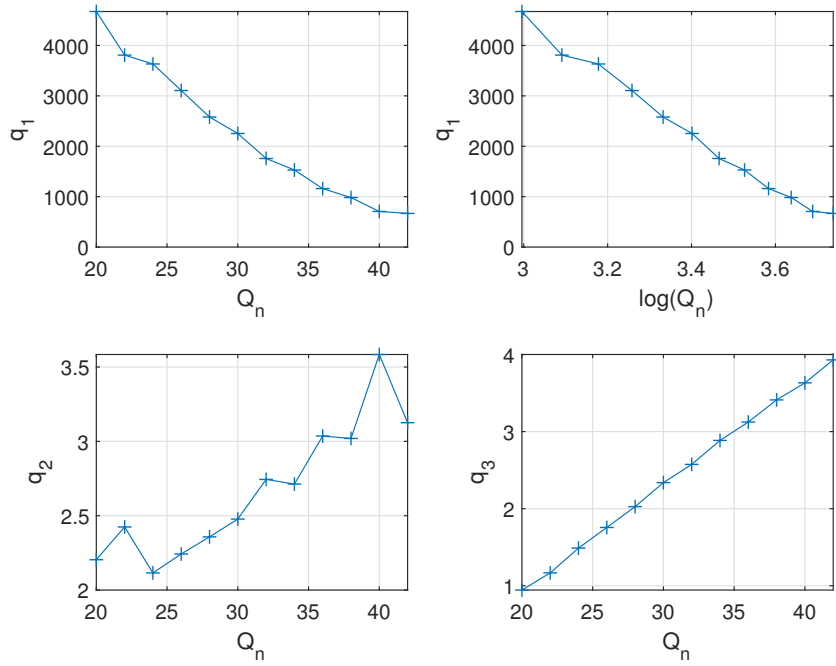


FIGURE 6 – Valeurs des estimées des paramètres q_1 , q_2 et q_3 obtenues pour différentes valeurs de $Q_n^{(i)}$, $i = 1, \dots, 12$ en fonction de Q_n (haut-gauche, bas-gauche, bas-droit) ou de $\log(Q_n)$ (haut-droit)

- (b) q_2 à Q_n ; elle sera notée $q_2 = g_2(Q_n, \mathbf{p})$;
- (c) q_3 à Q_n ; elle sera notée $q_3 = g_3(Q_n, \mathbf{p})$;
- 13. Donner la structure globale du modèle de comportement permettant de relier R_n , D_{n-1} et Q_n . Combien de paramètres ce modèle fait-il intervenir ?
- 14. Quel pourra être l'utilisation de ce modèle dans le contexte applicatif considéré ?

5 Estimation des paramètres d'un oscillateur

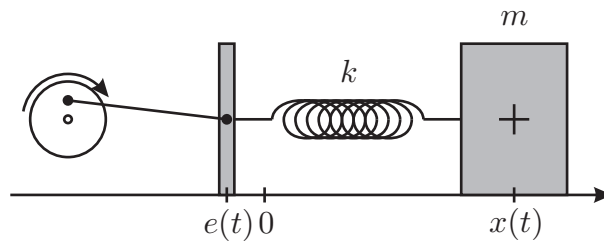


FIGURE 7 – Système masse-ressort

On considère le système masse-ressort représenté sur la figure 7. Un objet de masse m est relié à un support mobile par l'intermédiaire d'un ressort de constante de raideur k . A l'équilibre, le centre de gravité de l'objet se trouve en $x = \ell$ et la longueur du ressort est ℓ . A l'instant $t = 0$, l'objet est immobile et le support mobile est mis en mouvement par l'intermédiaire d'un cylindre tournant auquel il est relié par une bielle. La position du support mobile en fonction du temps est notée $e(t)$ et on suppose que $e(0) = 0$. On suppose que le mouvement est rectiligne et que l'objet subit un frottement

fluide dont la direction est opposée à la vitesse de déplacement et dont la norme est $\mu \left| \frac{dx}{dt} \right|$, où μ est la constante de frottement.

1. Donner l'équation différentielle du deuxième ordre reliant $y(t) = x(t) - \ell$ et $e(t)$ en fonction de m , μ et k . Préciser les conditions initiales.
2. En posant $y_1 = y$ et $y_2 = \frac{dy}{dt}$, donner le système différentiel du premier ordre décrivant l'évolution de $\mathbf{y} = (y_1, y_2)^T$. Préciser les conditions initiales.

On cherche à concevoir une expérience permettant d'estimer simultanément μ et k (m est mesuré à l'aide d'une balance). Pour cela, des mesures bruitées de la position du mobile $z(t_k) = y_1(t_k) + b(t_k)$ sont prélevées aux instants de mesures t_k , $k = 1, \dots, N$. L'identification du vecteur de paramètres $\mathbf{p} = (\mu, k)^T$ doit se faire sans résoudre à la main le système d'équations différentielles. Une procédure d'intégration numérique simple (par exemple une méthode d'Euler) sera employée pour simuler le modèle du système. L'estimation est réalisée par la méthode des moindres carrés. L'estimée de \mathbf{p} est l'argument du minimum du critère

$$c(\mathbf{p}) = \frac{1}{2} \sum_{k=1}^N (y_m(t_k, \mathbf{p}) - z(t_k))^2. \quad (4)$$

Nous supposons que $y_m(t, \mathbf{p}) = y_1(t)$.

1. Justifiez la forme du critère utilisé pour réaliser l'estimation de \mathbf{p} .
2. Montrez que le gradient du critère $c(\mathbf{p})$ s'exprime de la manière suivante

$$\mathbf{g}(\mathbf{p}) = \begin{pmatrix} \frac{\partial c(\mathbf{p})}{\partial p_1} \\ \frac{\partial c(\mathbf{p})}{\partial p_2} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^N (y_m(t_k) - z(t_k)) s_1^{p_1}(t_k) \\ \sum_{k=1}^N (y_m(t_k) - z(t_k)) s_1^{p_2}(t_k) \end{pmatrix} \quad (5)$$

où

$$s_i^{p_j}(t) = \frac{\partial y_i}{\partial p_j}(t)$$

sont les fonctions de sensibilité des y_i , $i = 1, 2$, à l'égard de p_j , $j = 1, 2$.

3. Donner un système d'équations différentielles dont la solution donne l'évolution de $s_1^{p_1}$ et $s_2^{p_1}$ en fonction du temps. Préciser les conditions initiales.
4. En déduire le système d'équations différentielles satisfaits par $s_1^{p_2}$ et $s_2^{p_2}$.
5. Les deux systèmes décrivant l'évolution des fonctions de sensibilité peuvent être regroupés au sein d'un seul grand système de six équations différentielles du premier ordre. Ecrire un fichier `ressort.m` permettant d'obtenir les dérivées temporelles des différentes fonctions de sensibilité en plus de dy_1/dt et dy_2/dt . La syntaxe de cette fonction devra être

```
% Entrées :
% - y : valeur de y à l'instant t
% - p : vecteur des paramètres
% - e : position du support à l'instant t
%
dy1dt =
dy2dt =
% Fonctions de sensibilité par rapport à p1
ds11dt =
ds21dt =
% Fonctions de sensibilité par rapport à p2
```

```

ds12dt =
ds22dt =
dydt = [dy1dt;dy2dt;ds11dt;ds21dt;ds12dt;ds22dt];

```

La fonction `euler.m`, dont l'en-tête est

```

function X = euler(func,x0,p,t,M)

%%
% Intégration par la méthode d'Euler de l'EDO décrite par func
% Entrées :
% - func : fonction à intégrer
% - x0 : condition initiales à t=0
% - p : vecteur des paramètres
% - t : vecteur des instants de mesure
% - M : nombre de pas d'intégration entre deux instants de mesure
%
% Sorties :
% - X : mesures mises sous forme de vecteurs colonnes rangés en une matrice

```

permet de réaliser une intégration numérique approchée d'un système d'équations différentielles ordinaires, partant d'un vecteur de conditions initiales, pour différents instants stockés dans un vecteur `t`, et avec une précision dépendant du nombre de pas de calcul M ($= 100$ par exemple) entre chaque instant de mesure.

Construire une fonction prenant en entrée `p`, `z` et `t` permettant de calculer la valeur ainsi que le gradient du critère (4). Une structure possible est

```
function [c,g]=critgrad(p,z,t)
```

Cette fonction utilisera les fonctions `euler.m` et `ressort.m`.

1. Rappelez la structure de la méthode d'optimisation du gradient.