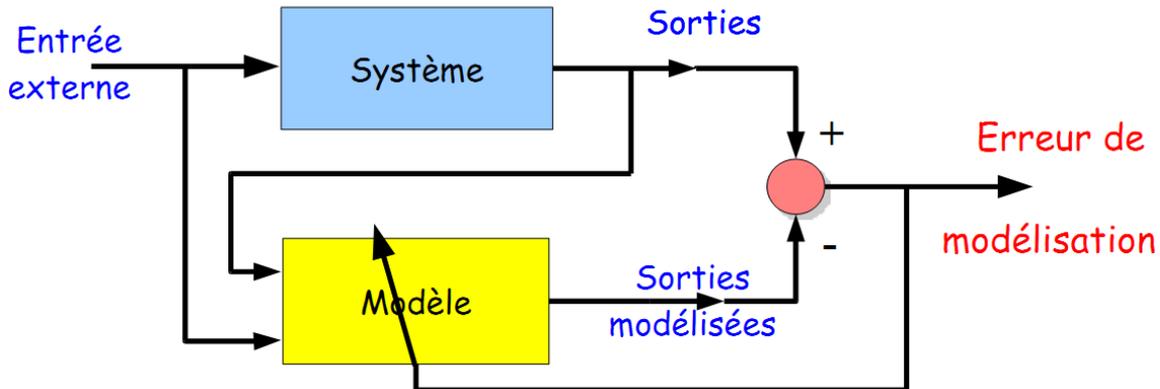


La détermination des paramètres d'un système à partir de l'observation de sa sortie, compte-tenu de l'entrée appliquée.



Le problème d'identification paramétrique peut être étudié dans un cadre plus général avec la méthode de l'erreur de prédiction. Cette méthode est basée sur les trois étapes suivantes:

1. Choisir la structure du modèle du système. Le choix de la structure dépend des hypothèses sur l'ordre du modèle du processus et la nature du bruit. Deux hypothèses différentes pour le bruit sont considérées:
 - (a) Le bruit n'est pas corrélé avec l'entrée du processus. Cette hypothèse conduit aux structures sans modèle du bruit.
 - (b) Le bruit sur la sortie est un bruit blanc filtré par un filtre d'ordre fini. Cette hypothèse conduit aux structures avec modèle du bruit.
2. Définir le prédicteur de la sortie en relation avec la structure du modèle. Le prédicteur est une fonction des paramètres du modèle à identifier (θ inconnu) et des signaux d'entrée et sortie, $\hat{y}(k) = \mathcal{F}(\theta, y(k-1), \dots, u(k-1), \dots)$. En
3. Minimiser l'erreur de prédiction avec une méthode numérique. Après avoir défini le prédicteur pour la structure choisie, on obtient l'erreur de prédiction $\varepsilon(k) = y(k) - \hat{y}(k)$ qui est une fonction du vecteur des paramètres à identifier θ . On obtient ainsi un problème d'optimisation numérique dont le critère à minimiser est la norme 2 de l'erreur de prédiction:

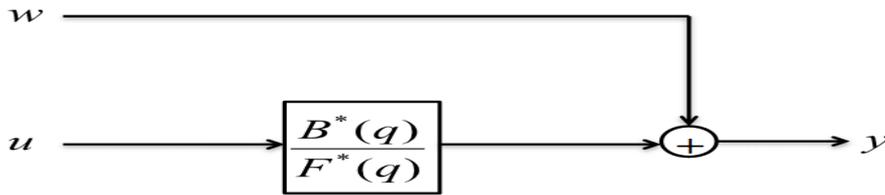
$$\hat{\theta} = \arg \min_{\theta} J(\theta) = \frac{1}{N} \sum_{k=1}^N \varepsilon^2(k) \quad (3.1)$$

Structures sans modèle du bruit

Modèle Erreur de sortie (« Output Error »)

$$y[k] = \frac{B^*(q)}{F^*(q)} u[k] + w[k]$$

$$F^*(q) = 1 + f_1^* q^{-1} + \dots + f_{n_f}^* q^{-n_f}$$



$$\theta^* = [b_1^* \quad \dots \quad b_{n_b}^* \quad f_1^* \quad \dots \quad f_{n_f}^*]^T$$

Modèle de prédiction

$$\hat{y}[k, \theta] = \frac{B(q)}{F(q)} u[k] \quad \text{Modèle non LRP}$$

$$e[k, \theta] = y[k] - \hat{y}[k, \theta] = y[k] - \frac{B(q)}{F(q)} u[k]$$

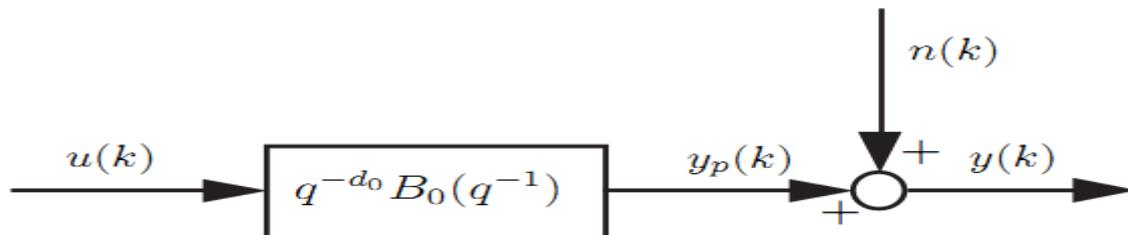
$$= w[k] \quad \text{si} \quad \theta = \theta^*$$

$$\hat{y}[k, \theta] = -f_1 \hat{y}[k-1, \theta] - \dots - f_{n_f} \hat{y}[k-n_f, \theta] + b_1 u[k-n_r] + \dots + b_{n_b} u[k-n_r-n_b+1]$$

$$= \varphi^T[k-1, \theta] \theta \quad \text{Modèle pseudo-LRP de simulation}$$

Structure FIR

Un cas particulier de la structure OE est le cas où $n_0 = 0$ ou $A_0(q^{-1}) = 1$. Ceci est nommé la structure de FIR (finite impulse response ou réponse impulsionnelle finie) (fig. 3.10).



Pour cette structure on a:

$$y(k) = q^{-d} B_0(q^{-1}) u(k) + n(k)$$

La prédiction de la sortie de ce modèle s'écrit:

$$\hat{y}(k) = q^{-d} B(q^{-1}) u(k)$$

On calcule ensuite l'erreur de prédiction:

$$\varepsilon(k) = y(k) - \hat{y}(k) = y(k) - \varphi_f^T(k)\theta$$

où

$$\begin{aligned}\varphi_f^T(k) &= [u(k-d), u(k-d-1), \dots, u(k-d-m)] \\ \theta^T &= [b_0, b_1, \dots, b_m]\end{aligned}$$

Structures avec modèle du bruit

A. Modèle ARX (AutoRégressif à variable eXogène)

$$\begin{aligned}y[k] + a_1^* y[k-1] + \dots + a_{n_a}^* y[k-n_a] \\ = b_1^* u[k-n_r] + b_2^* u[k-n_r-1] + \dots + b_{n_b}^* u[k-n_r-n_b+1] + w[k]\end{aligned}$$

Bruit blanc de moyenne nulle 

$y[k] = y(kT)$ où $T =$ Période d'échantillonnage

* Signifie « vraie » valeur des paramètres

n_r (≥ 1 en général) Retard pur du système

$\{n_a, n_b, n_r\} =$ Structure de modèle

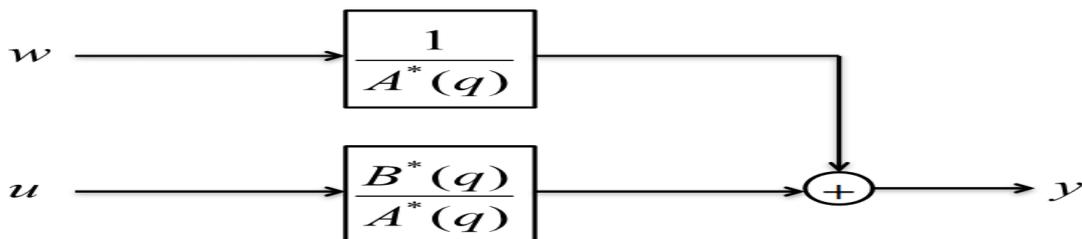
$\theta^* = [a_1^* \dots a_{n_a}^* \quad b_1^* \quad b_2^* \dots b_{n_b}^*]^T$ Vecteur des « vraies » valeurs du modèle

Représentation polynomiale

$$A^*(q) = 1 + a_1^* q^{-1} + a_2^* q^{-2} + \dots + a_{n_a}^* q^{-n_a}$$

$$B^*(q) = b_1^* q^{-n_r} + b_2^* q^{-n_r-1} + \dots + b_{n_b}^* q^{-n_r-n_b+1}$$

$$A^*(q)y[k] = B^*(q)u[k] + w[k]$$



Modèle de prédiction

$\theta = [a_1 \quad a_2 \quad \dots \quad a_{n_a} \quad b_1 \quad b_2 \quad \dots \quad b_{n_b}]^T$ Estimée des paramètres

$$\hat{y}[k, \theta] = -a_1 y[k-1] - \dots - a_{n_a} y[k-n_a] + b_1 u[k-n_r] + \dots + b_{n_b} u[k-n_r-n_b+1]$$

Prédiction de la sortie à l'instant k

$$e[k, \theta] = y[k] - \hat{y}[k, \theta] \quad \text{Erreur de prédiction (à 1 pas)}$$

$$= w[k] \quad \text{si} \quad \theta = \theta^*$$

$$\Rightarrow E\{e[k, \theta^*]\} = 0$$

Forme polynomiale:

$$\hat{y}[k, \theta] = (1 - A(q))y[k] + B(q)u[k]$$

$$= \varphi^T[k-1]\theta \quad \text{Modèle LRP}$$

$$\varphi^T[k-1] = [-y[k-1] \quad \dots \quad -y[k-n_a] \quad u[k-n_r] \quad \dots \quad u[k-n_r-n_b+1]]$$

Vecteur régresseur

$$e[k, \theta] = y[k] - \hat{y}[k, \theta] = \varphi^T[k-1](\theta^* - \theta) + w[k]$$

B. Modèle ARMAX (AutoRégressif à Moyenne Ajustée et variable eXogène)

$$y[k] + a_1^* y[k-1] + \dots + a_{n_a}^* y[k-n_a] = b_1^* u[k-n_r] + \dots + b_{n_b}^* u[k-n_r-n_b+1]$$

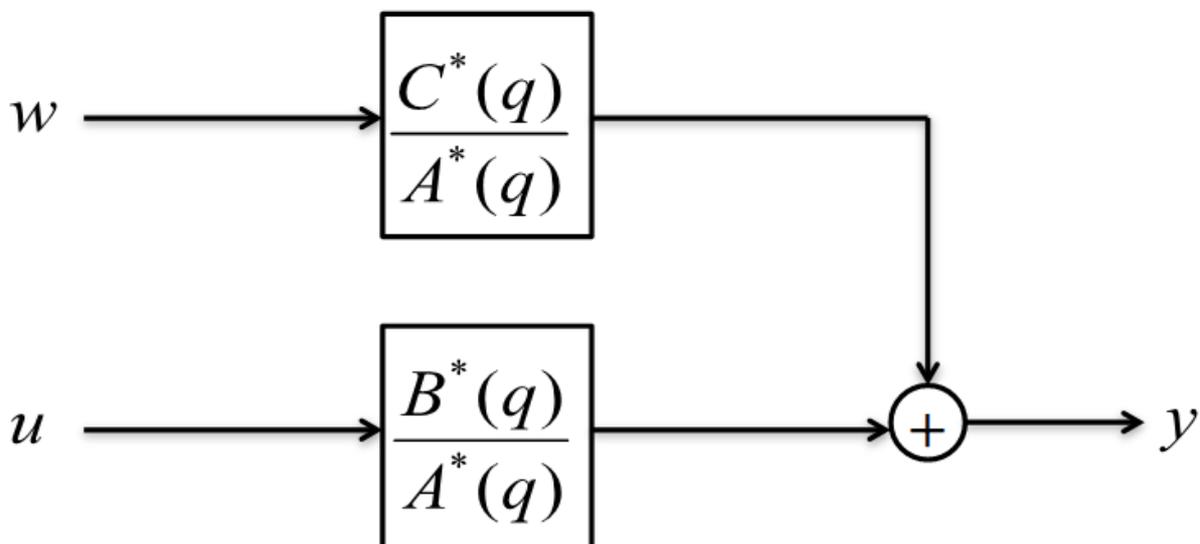
$$+ w[k] + c_1^* w[k-1] + \dots + c_{n_c}^* w[k-n_c]$$

Bruit blanc de moyenne nulle ↗

$$\theta^* = [a_1^* \quad \dots \quad a_{n_a}^* \quad b_1^* \quad \dots \quad b_{n_b}^* \quad c_1^* \quad \dots \quad c_{n_c}^*]^T$$

$$C^*(q) = 1 + c_1^* q^{-1} + c_2^* q^{-2} + \dots + c_{n_c}^* q^{-n_c}$$

$$A^*(q)y[k] = B^*(q)u[k] + C^*(q)w[k]$$



Modèle de prédiction

$$y[k] = \left(1 - \frac{A^*(q)}{C^*(q)}\right) y[k] + \frac{B^*(q)}{C^*(q)} u[k] + w[k]$$

$$\hat{y}[k, \theta] = \frac{B(q)}{C(q)} u[k] + \left(\frac{C(q) - A(q)}{C(q)}\right) y[k]$$

Prédiction à un pas

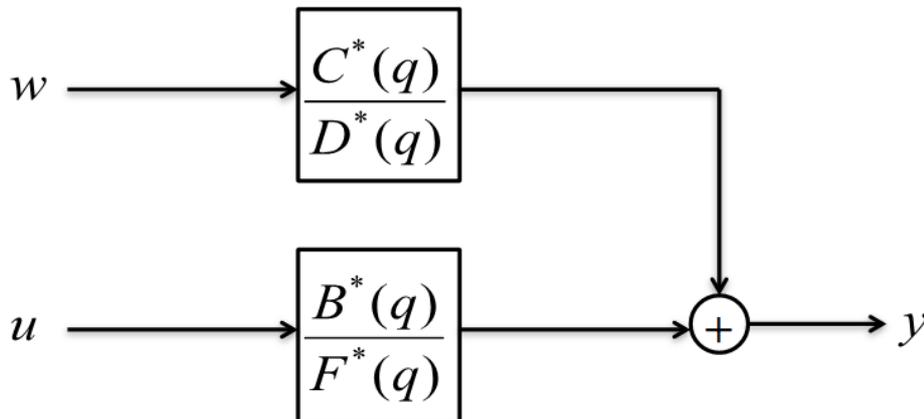
Modèle non LRP

$$e[k, \theta] = y[k] - \hat{y}[k, \theta] = -\frac{B(q)}{C(q)} u[k] + \frac{A(q)}{C(q)} y[k] \quad \text{Erreur de prédiction}$$

$$= w[k] \quad \text{si} \quad \theta = \theta^*$$

Modèle de BOX et JENKINS

$$y[k] = \frac{B^*(q)}{F^*(q)} u[k] + \frac{C^*(q)}{D^*(q)} w[k]$$



Modèle de prédiction

$$\hat{y}[k, \theta] = \frac{D(q)B(q)}{C(q)F(q)} u[k] + \left(1 - \frac{D(q)}{C(q)}\right) y[k]$$

$$e[k, \theta] = y[k] - \hat{y}[k, \theta]$$

$$= w[k] \quad \text{si} \quad \theta = \theta^*$$

Minimisation de l'erreur de prédiction

On a vu que les paramètres d'un modèle sont identifiés en minimisant un critère quadratique basé sur l'erreur de prédiction définie comme:

$$J(\theta) = \frac{1}{N} \sum_{k=1}^N \varepsilon^2(k) = \frac{1}{N} \sum_{k=1}^N [y(k) - \hat{y}(k)]^2$$

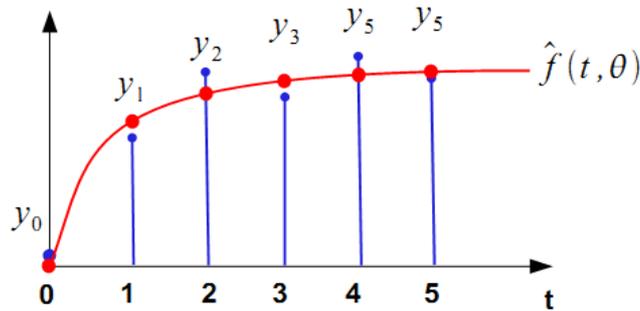
Pour les structures FIR et ARX, l'erreur de prédiction est linéaire par rapport au vecteur de paramètres à identifier et donc le minimum du critère () peut être calculé avec la méthode des moindres carrés (). Les autres structures résultent en un problème d'optimisation non linéaire pour laquelle on utilise le plus souvent l'algorithme itératif de Gauss-Newton.

● Méthode de Gauss-Newton (méthode de gradient)

● Position du problème

Dans de nombreuses applications, la fonction modèle n'est pas linéaire par rapport à ses paramètres.

Exemple : réponse indicielle d'un circuit du 1er ordre



L'expression du modèle est : $f(t, \theta) = K(1 - e^{-\frac{t}{\tau}})$

avec $\theta = [K, \tau]^T$

⇒ la dépendance vis à vis du paramètre τ est non-linéaire, il n'est plus possible de résoudre le système par la méthode linéaire (pas de matrice H)

⇒ La proposition est de partir d'une valeur initiale θ_0 des paramètres et de modifier itérativement la valeur de θ d'un incrément δ_θ de façon à minimiser le critère d'erreur quadratique cumulée \mathcal{E} à chaque étape.

Etape initiale

Pour $\theta_0 = [a_{10}, a_{20}, \dots]^T$ le modèle prend les valeurs : $F(t, \theta_0) = \begin{bmatrix} f(t_1, \theta_0) \\ f(t_2, \theta_0) \\ \dots \\ f(t_n, \theta_0) \end{bmatrix}$

L'erreur entre le modèle et les mesures est : $\underline{\varepsilon} = \underline{Y} - \underline{F}(t, \theta_0)$

L'erreur quadratique cumulée est : $\mathcal{E}(\theta_0) = \underline{\varepsilon}^T \underline{\varepsilon} = (\underline{Y} - \underline{F}(t, \theta_0))^T (\underline{Y} - \underline{F}(t, \theta_0))$

Généralement, l'erreur cumulée sera importante, les conditions initiales étant éloignées de la solution optimale.

Incrémentation

⇒ modification de la valeur de θ_0 d'un incrément δ_θ de façon à minimiser le critère

d'erreur quadratique cumulée \mathcal{E}

Pour $\theta = \theta_0 + \delta_\theta$, le modèle prend les valeurs $F(t, \theta_0 + \delta_\theta)$ et l'erreur $\mathcal{E}(\theta_0 + \delta_\theta)$

⇒ Pour se placer au minimum d'erreur, on choisit δ_θ tel que $\frac{\partial \mathcal{E}(\theta_0 + \delta_\theta)}{\partial \delta_\theta} = 0$

Calcul du minimum de l'erreur cumulée :

$$\frac{\partial \mathcal{E}(\theta_0 + \delta_\theta)}{\partial \delta_\theta} = -2 \left[\frac{\partial F(t, \theta_0 + \delta_\theta)}{\partial \delta_\theta} \right]^T (\underline{Y} - \underline{F}(t, \theta_0 + \delta_\theta)) = 0$$

Calcul de $F(t, \theta_0 + \delta_\theta)$

Le développement de Taylor du 1er ordre du modèle permet d'approximer la nouvelle valeur du modèle à chaque instant d'observation :

$$f(t_i, \theta_0 + \delta_\theta) = f(t_i, \theta_0) + J_i(\theta_0) \cdot \delta_\theta$$

avec $J_i(\theta_0) = \nabla (f(t_i, \theta_0))^T$ gradient de f en ligne

L'extension à l'ensemble des points de calculs prend la forme matricielle suivante :

$$F(t, \theta_0 + \delta) = F(t, \theta_0) + J(\theta_0) \cdot \delta_\theta$$

$$\text{Avec : } J(\theta_0) = \frac{\partial F(t, \theta_0 + \delta_\theta)}{\partial \delta_\theta} = \begin{bmatrix} \frac{\partial f(t_1)}{\partial a_1} & \frac{\partial f(t_1)}{\partial a_2} & \dots & \frac{\partial f(t_1)}{\partial a_k} \\ \frac{\partial f(t_2)}{\partial a_1} & \frac{\partial f(t_2)}{\partial a_2} & \dots & \frac{\partial f(t_2)}{\partial a_k} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f(t_n)}{\partial a_1} & \frac{\partial f(t_n)}{\partial a_2} & \dots & \frac{\partial f(t_n)}{\partial a_k} \end{bmatrix}$$

jacobienne de f /paramètres

Calcul du minimum de l'erreur cumulée :

$$\frac{\partial \mathcal{E}(\theta_0 + \delta_\theta)}{\partial \delta_\theta} = -2 \left[\frac{\partial F(t, \theta_0 + \delta_\theta)}{\partial \delta_\theta} \right]^T (\underline{Y} - \underline{F}(t, \theta_0 + \delta_\theta)) = 0$$

D'où : $-2 J^T (\underline{Y} - (\underline{F}(t, \theta_0) + J(\theta_0) \cdot \delta_\theta)) = 0$

On en déduit la valeur de l'accroissement à faire sur les paramètres pour minimiser l'erreur:

$$\delta_\theta = [J(\theta_0)^T \cdot J(\theta_0)]^{-1} \cdot J(\theta_0)^T (\underline{Y} - \underline{F}(t, \theta_0))$$

Itération

On itère en définissant les nouvelles valeurs des paramètres $\theta_1 = \theta_0 + \delta_\theta$ et les nouvelles valeurs du modèle $\underline{F}(t, \theta_1)$. La correction suivante à faire sera :

$$\delta_\theta = [J(\theta_1)^T \cdot J(\theta_1)]^{-1} \cdot J(\theta_1)^T (\underline{Y} - \underline{F}(t, \theta_1))$$

Attention : certaines valeurs initiales ne permettent pas à l'algorithme de converger

Limitation : l'inversion de la matrice $J^T J$ peut poser problème (matrice singulière).

Pour éviter ce blocage, l'algorithme a été modifié par **Levenberg-Marquardt**:

$$\delta_\theta = [J^T \cdot J + \lambda I]^{-1} \cdot J^T (\underline{Y} - \underline{F}(t, \theta))$$

Le paramètre λ joue le rôle d'un amortissement de la correction; il doit être ajusté à chaque pas de calcul. Dans les cas simples, on peut se contenter d'un amortissement constant, dont la valeur initiale a été proposée par Marquardt:

$$\lambda_0 = \tau \cdot \max[J^T J]$$

τ est un paramètre de gain à choisir convenablement (!)

Exemple d'application : soit un système du 1er ordre dont désire connaître les paramètres caractéristiques (gain et constante de temps) par l'observation de la réponse indicielle.

Temps en s	0	1	2	3	4	5
y(t)	0.05	0.45	0.59	0.64	0.64	0.69

L'observation directe montre que le gain (*valeur finale*) est proche de 0.7 et la constante de temps de l'ordre de la seconde (*échelle de temps*)

Le modèle de la réponse indicielle est : $f(t, \theta) = K(1 - e^{-\frac{t}{\tau}})$ avec $\theta = [K, \tau]^T$

La matrice du Jacobien est construite à partir des dérivées partielles de f par rapport à θ

$$\frac{\partial f(t, \theta)}{\partial K} = 1 - e^{-\frac{t}{\tau}} \quad \frac{\partial f(t, \theta)}{\partial \tau} = -K \frac{t}{\tau^2} e^{-\frac{t}{\tau}}$$

Contrairement au cas d'un modèle linéaire/paramètres, les dérivées partielles dépendent des paramètres eux-mêmes.

Il faut fixer une valeur initiale de K et θ pour donner une valeur à la matrice Jacobienne

Valeurs initiales proposées : $K_0 = 1 \quad \tau_0 = 1$

- méthode de Gauss-Newton :

La matrice Jacobienne est : $J(\theta_0) = \begin{bmatrix} 0. & 0. \\ 0.63 & -0.37 \\ 0.86 & -0.27 \\ 0.95 & -0.15 \\ 0.98 & -0.07 \\ 0.99 & -0.03 \end{bmatrix}$ les valeurs du modèle : $F(\theta_0) = \begin{bmatrix} 0. \\ 0.63 \\ 0.86 \\ 0.95 \\ 0.98 \\ 0.99 \end{bmatrix}$

L'erreur quadratique cumulée est de : $\mathcal{E} = 0.42$

L'incrément à faire sur les paramètres est : $\delta_\theta = (J^T J)^{-1} J^T (Y - F(\theta_0)) = \begin{bmatrix} -0.33 \\ -0.06 \end{bmatrix}$

La nouvelle valeur des paramètres est : $\theta_1 = \begin{bmatrix} 0.67 \\ 0.94 \end{bmatrix}$

Après 10 itérations , la valeur des paramètres converge vers : $\theta_{10} = \begin{bmatrix} 0.67 \\ 0.92 \end{bmatrix}$

L'erreur quadratique cumulée est de : $\mathcal{E} = 0.0035$