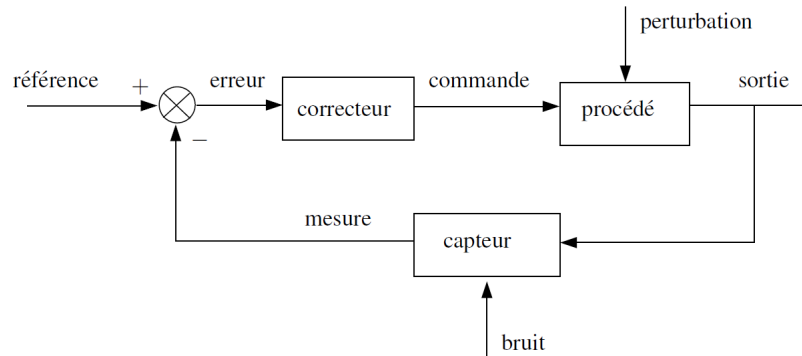


Automatique



Ce document regroupe les principaux résultats applicables pour la commande et l'estimation des systèmes dynamiques linéaires à plusieurs entrées et plusieurs sorties, en utilisant le formalisme de la représentation d'état. Les principales variables impliquées sont :

- le vecteur d'entrée (ou de commande) \mathbf{u} , de dimension m
- le vecteur d'état \mathbf{x} , de dimension n
- le vecteur de sortie (ou de mesure) \mathbf{y} , de dimension p

1 Fonction de transfert

1.1 Systèmes LTI

Les systèmes linéaires à temps continu invariants dans le temps ont les propriétés suivantes :

Linéarité Soit y_1 et y_2 sont les réponses du système Σ excité séparément par les entrées u_1 et u_2 respectivement. Soit α un réel quelconque. Le système est linéaire si sa sortie vaut $\alpha \cdot y_1 + y_2$ en réponse à l'entrée $\alpha \cdot u_1 + u_2$.

Invariance temporelle Un système est dit invariant si une même commande, appliquée à deux instants différents produit la même sortie.

Causalité Un système d'entrée u et de sortie y est causal si, $\forall t < 0, u = 0 \implies y = 0$. Autrement dit, la réponse du système ne précède pas son excitation.

1.2 Fonction de transfert

1.2.1 Transformée de Laplace

La réponse d'un système LTI d'entrée u et de sortie y est le produit de convolution : $y = g * u$ où g est la réponse impulsionnelle du système, obtenue en excitant le système par une impulsion de Dirac (de durée nulle et d'amplitude infinie).

La transformée de Laplace est définie par : $F(s) = \mathcal{L}\{f(t)\} = \int_0^{+\infty} f(t) \exp(-st) dt$, où s est la variable de Laplace. Elle transforme le produit de convolution en un produit, et donc les équations différentielles en équations algébriques. Notamment, la dérivation devient un produit par s et l'intégration un quotient en $\frac{1}{s}$.

En appliquant Laplace au produit de convolution, on obtient la fonction de transfert :

$$G(s) = \frac{Y(s)}{U(s)} = \frac{\sum_{i=0}^m b_i s^i}{\sum_{i=c}^n a_i s^i} = \frac{N(s)}{D(s)}$$

Les racines du numérateur N sont les zéros du système et les racines du dénominateur D en sont les pôles.

1.2.2 Gain et phase

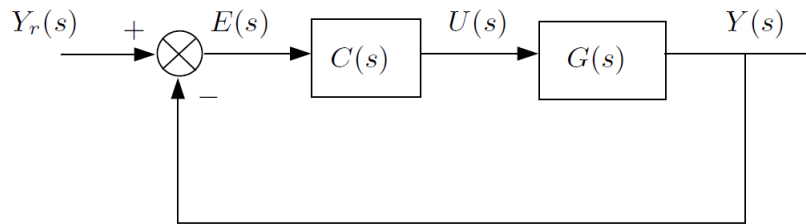
La réponse harmonique du système ayant pour fonction de transfert $G(s)$, avec $s = j\omega$ (ω étant la pulsation du régime permanent sinusoïdal), est la réponse à l'entrée sinusoïdale $u(t) = A \sin(\omega t)$ telle que

$$y(t) = A \|G(j\omega)\| \sin(\omega t + \arg\{G(j\omega)\})$$

Le comportement fréquentiel du système se caractérise par le module et l'argument de $G(j\omega)$:

$$\begin{cases} G_{dB}(\omega) = 20 \log \|G(j\omega)\| \\ \varphi(\omega) = \arg\{G(j\omega)\} \end{cases}$$

1.3 Boucle ouverte et boucle fermée



On note $C(s)$ le correcteur à calculer et $G(s)$ la fonction de transfert du système. La fonction de transfert en boucle ouverte est la $FTBO = C(s)G(s)$. La boucle fermée (transfert entre la référence Y_r et la sortie Y) est :

$$FTBF = \frac{FTBO}{1 + FTBO} = \frac{CG}{1 + CG}$$

Un système LTI (en BO ou en BF) est stable si tous ses pôles sont à partie réelle strictement négative. On définit :

- la marge de phase $M_\varphi = 180^\circ + \arg\{CG(j\omega_c)\}$ avec ω_c tel que $\|CG(j\omega_c)\| = 0$
- la marge de gain $M_G = -\|CG(j\omega_\pi)\|$ avec ω_π tel que $\arg\{CG(j\omega_\pi)\} = -180^\circ$

1.4 Correcteurs linéaires (PID)

Nom	Expression	Avantages	Inconvénients
Proportionnel	K	augmente le gain en basse fréquence	risque de perte de robustesse (la marge de phase diminue)
Intégral	$\frac{K}{s}$	augmente le gain, assure le rejet asymptotique des perturbations constantes	risque de déstabilisation
Avance de phase (PD)	$K \cdot \frac{1+as}{1+Ts}$	augmente la phase, se prête bien à la correction des systèmes peu stables	attention au choix de a ($1 < a < 10$ en pratique) et T
Proportionnel Intégral (PI)	$\frac{1+KTs}{Ts}$	augmente le gain, assure le rejet de perturbations, conserve la phase en haute fréquence	Introduit un déphasage qui peut être gênant pour la marge de phase
PID	$K \cdot \left(1 + \frac{1}{T_i s} + T_p s\right)$	réalise toutes les actions précédentes	réglage pas toujours aisé

2 Systèmes multivariables à temps continu

2.1 Forme d'état

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \end{cases}$$

- \mathbf{A} est la matrice d'état, de dimension $n \times n$.
- \mathbf{B} est la matrice d'entrée, de dimension $n \times m$.
- \mathbf{C} est la matrice de sortie, de dimension $p \times n$.
- \mathbf{D} est la matrice directe entrée/sortie (généralement nulle), de dimension $p \times m$.

La solution exacte de l'équation différentielle $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$ dans le domaine temporel est donnée par :

$$\mathbf{x}(t) = \exp(\mathbf{A}(t - t_0)) \mathbf{x}(t_0) + \int_{t_0}^t \exp(\mathbf{A}(t - \tau)) \mathbf{B}\mathbf{u}(\tau) d\tau$$

2.2 Matrice de transfert

$$\mathbf{y} = [\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}] \mathbf{u}$$

$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$ est la matrice de transfert du système.

2.3 Stabilité

Un système linéaire continu est asymptotiquement stable si les valeurs propres de la matrice d'état \mathbf{A} sont à parties réelles strictement négatives. Les valeurs propres de la matrice d'état sont les pôles de la matrice de transfert. Plus la partie réelle des pôles est grande, plus la réponse est rapide ; plus la partie imaginaire des pôles est grande, plus le système est oscillant.

2.4 Commandabilité

La propriété de commandabilité vérifie qu'à partir des entrées \mathbf{u} , il est possible de modifier *toutes* les variables d'état \mathbf{x} . Un système $\{\mathbf{A}, \mathbf{B}\}$ est commandable si et seulement si la matrice de commandabilité $\mathcal{C}(\mathbf{A}, \mathbf{B}) = [\mathbf{B}, \mathbf{A}\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}]$ est de rang n .

2.5 Commande par retour d'état

Le fait que le système soit commandable assure qu'il est possible, par le choix d'une matrice \mathbf{K} appropriée, de placer arbitrairement les valeurs propres du système en boucle fermée avec la commande par retour d'état $\mathbf{u} = -\mathbf{K} \cdot \mathbf{x}$. Les pôles du système en boucle fermée sont les valeurs propres de $\mathbf{A} - \mathbf{B}\mathbf{K}$.

2.6 Observabilité

L'observabilité vérifie la possibilité de reconstruire la valeur de l'état du système \mathbf{x} à partir de la connaissance des entrées \mathbf{u} et des sorties \mathbf{y} . Le système $\{\mathbf{A}, \mathbf{C}\}$ est observable si et seulement si la matrice d'observabilité

$$\mathcal{O}(\mathbf{A}, \mathbf{C}) = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \vdots \\ \mathbf{C}\mathbf{A}^{n-1} \end{bmatrix} \text{ est de rang } n.$$

2.7 Observateur de Luenberger

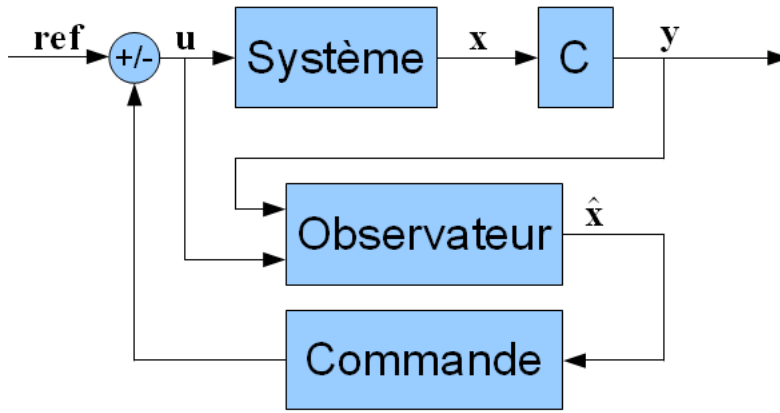
Le but de l'observateur est d'obtenir une estimée de l'état \mathbf{x} , notée $\hat{\mathbf{x}}$. L'observateur utilise les informations structurelles présentes dans les matrices \mathbf{A} , \mathbf{B} et \mathbf{C} pour "prédire" à l'aide de l'équation dynamique d'état la future valeur de \mathbf{x} . La mesure \mathbf{y} fournie par les capteurs est utilisée pour "corriger" la précédente prédiction et ainsi fournir l'estimation la plus fiable possible de l'état. La formulation de l'observateur est la suivante :

$$\begin{cases} \dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{L}(\mathbf{y} - \hat{\mathbf{y}}) \\ \hat{\mathbf{y}} = \mathbf{C}\hat{\mathbf{x}} \end{cases}$$

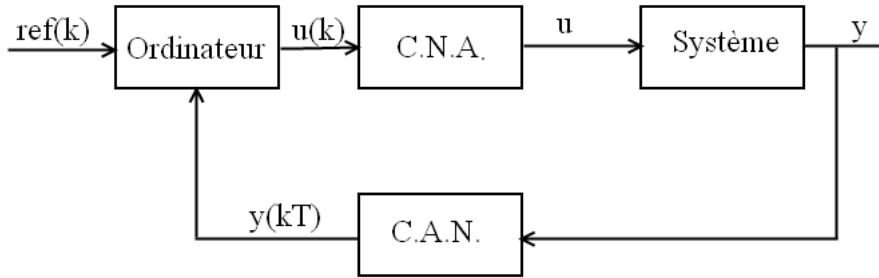
L'écart entre l'estimation $\hat{\mathbf{x}}$ et la valeur vraie (inconnue) \mathbf{x} se calcule comme suit :

$$\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}} \Rightarrow \dot{\mathbf{e}} = (\mathbf{A} - \mathbf{L}\mathbf{C})\mathbf{e}$$

Pour que l'erreur d'estimation tende vers 0, il faut que $\mathbf{A} - \mathbf{L}\mathbf{C}$ soit stable. Si le système est observable, la matrice de gain de l'observateur \mathbf{L} peut être choisie pour placer arbitrairement les valeurs propres de $\mathbf{A} - \mathbf{L}\mathbf{C}$.



3 Systèmes en temps discret



3.1 Forme d'état

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \end{cases}$$

La discrétisation exacte d'un système continu $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$ avec une période d'échantillonnage T est :

$$\mathbf{x}_{k+1} = \exp(\mathbf{A}T)\mathbf{x}_k + \left(\int_0^T \exp(\mathbf{A}t)dt \right) \mathbf{B}\mathbf{u}_k$$

Les méthodes approchées les plus usuelles pour la discrétisation sont :

Euler explicite : $\dot{x}(t) \approx \frac{x(t+T) - x(t)}{T}$; Euler implicite : $\dot{x}(t) \approx \frac{x(t) - x(t-T)}{T}$; Bilineaire : $\frac{\dot{x}(t) + \dot{x}(t+T)}{2} \approx \frac{x(t+T) - x(t)}{T}$.

3.2 Stabilité

Un système dynamique à temps discret est stable si les valeurs propres de la matrice d'état \mathbf{A} sont toutes de module strictement inférieur à 1 (cercle unité). Ce résultat se transpose également aux notions de commandabilité (matrice $\mathbf{A} - \mathbf{BK}$) et d'observabilité (matrice $\mathbf{A} - \mathbf{LC}$).

3.3 Filtre de Kalman

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{v}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{w}_k \end{cases}$$

On introduit dans le modèle les bruits d'état \mathbf{v}_k et de mesure \mathbf{w}_k , supposés Gaussiens de moyenne nulle. Leurs moyennes et covariances sont données par :

$$\begin{aligned} \mathcal{E}\{\mathbf{v}_k\} &= 0 \\ \mathcal{V}\{\mathbf{v}\} &= \mathbf{r} \\ \mathcal{E}\{\mathbf{w}_k\} &= 0 \\ \mathcal{V}\{\mathbf{w}\} &= \mathbf{q} \end{aligned}$$

La covariance de l'état estimé est notée \mathbf{P} , $\hat{\mathbf{x}}_{k|k-1}$ désigne la valeur de l'état prédite ou estimée à l'instant k à partir des informations disponibles jusqu'à l'instant $k-1$. Le filtre prédictor-estimateur optimal au sens de la minimisation de la variance de l'erreur d'estimation est donné par :

$$\begin{aligned} \hat{\mathbf{x}}_{k/k-1} &= \mathbf{A}\hat{\mathbf{x}}_{k-1/k-1} + \mathbf{B}\mathbf{u}_{k-1} \\ \mathbf{P}_{k/k-1} &= \mathbf{A}\mathbf{P}_{k-1/k-1}\mathbf{A}^T + \mathbf{r}_k \\ \hat{\mathbf{x}}_{k/k} &= \hat{\mathbf{x}}_{k/k-1} + \mathbf{L}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k) \\ \hat{\mathbf{y}}_k &= \mathbf{C}\hat{\mathbf{x}}_{k/k-1} \\ \mathbf{L}_k &= \mathbf{P}_{k/k-1}\mathbf{C}^T(\mathbf{C}\mathbf{P}_{k/k-1}\mathbf{C}^T + \mathbf{q}_k)^{-1} \\ \mathbf{P}_{k/k} &= (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k/k-1} \end{aligned}$$

3.4 Estimation de paramètres

Soit le modèle linéaire en les paramètres $\mathbf{y} = \mathbf{H}\mathbf{p} + \mathbf{w}$. \mathbf{H} est la matrice d'observation du système, connue, et \mathbf{w} est un bruit blanc gaussien centré. On veut donc trouver $\hat{\mathbf{p}}_{mc} = \arg \min \|\mathbf{y} - \mathbf{H}\mathbf{p}\|_2^2$.

Le gradient du critère des moindres carrés s'écrit :

$$\frac{\partial j}{\partial \mathbf{p}} = -2\mathbf{H}^T\mathbf{y} + 2\mathbf{H}^T\mathbf{H}\mathbf{p} = -2\mathbf{H}^T(\mathbf{y} - \mathbf{H}\mathbf{p}) = -2\mathbf{H}^T\mathbf{e}$$

Le minimum global est obtenu lorsque le gradient s'annule, soit lorsque :

$$\begin{aligned} -2\mathbf{H}^T\mathbf{y} + 2\mathbf{H}^T\mathbf{H}\hat{\mathbf{p}}_{mc} &= 0 \\ \Rightarrow \mathbf{H}^T\mathbf{H}\hat{\mathbf{p}}_{mc} &= \mathbf{H}^T\mathbf{y} \\ \Rightarrow \hat{\mathbf{p}}_{mc} &= (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y} \end{aligned}$$

Cette formule donne également la "pseudo-inverse" d'un système linéaire avec plus d'équations que d'inconnues.

4 Exemples de systèmes

4.1 Régulateur de vitesse

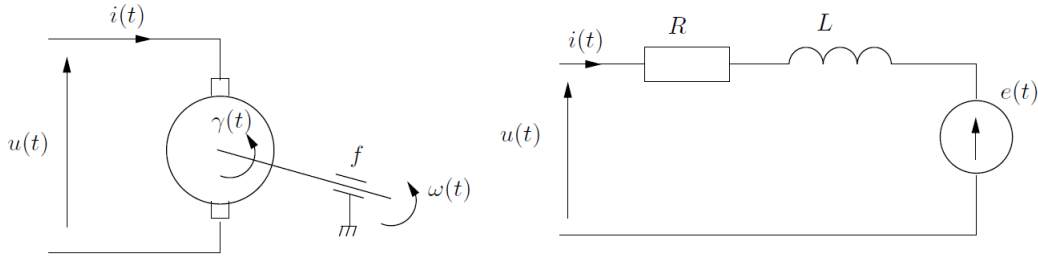
On considère la régulation d'une vitesse v à l'aide d'une commande d'accélération u . Le système dynamique est simplement défini comme un intégrateur :

$$\dot{v} = u$$

En temps discret (période T), le système devient :

$$v_{k+1} = v_k + Tu_k$$

4.2 Le moteur à courant continu



La modélisation la plus simple du moteur à courant continu est un circuit RLC. Les équations de la physique sont les suivantes :

$$\begin{cases} u = Ri + L \frac{di}{dt} + e \\ J \frac{d\omega}{dt} = \gamma - f\omega \end{cases}$$

où $e = K_e \omega$ et $\gamma = K_m i$. Dans cette modélisation linéaire, les pertes sont négligées.

Pour mettre ce système sous forme d'état, il faut tout d'abord choisir les variables d'état du système. La logique nous pousse à choisir les variables ayant une signification physique : $x_1 = i$ et $x_2 = \omega$, ce qui donne le vecteur d'état à deux dimensions $\mathbf{x} = [x_1, x_2]^T$. Le choix est également dicté par le fait que les dérivées de ces variables (et donc la dynamique) interviennent dans les équations. La commande \mathbf{u} est égale à la tension u . On mesure $y = \omega$. On réécrit alors le système sous la forme :

$$\begin{cases} L \frac{dx_1}{dt} = -Rx_1 - K_e x_2 + u \\ J \frac{dx_2}{dt} = K_m x_1 - f x_2 \end{cases}$$

4.3 Drone

Le modèle dynamique (légèrement simplifié) d'un drone multi-rotors est donné par :

$$\begin{cases} \dot{p} = v \\ m\dot{v} = \mathcal{T}Re_3 - mge_3 \\ \dot{\eta} = \Omega \\ \dot{\Omega} = J^{-1}\Gamma \end{cases}$$

avec :

- $p = [p_x, p_y, p_z]^T$ la position et $v = [v_x, v_y, v_z]^T$ la vitesse en repère inertiel (par rapport à une référence fixe), \mathcal{T} la poussée, m la masse du drone.
- R la matrice de rotation permettant de passer du repère inertiel au repère lié au drone avec $\eta = [\phi, \theta, \psi]^T$ les trois angles en roulis, tangage et lacet. e_3 est l'axe vertical du repère inertiel (où la gravité s'applique). Avec la convention de rotation ZYX ($\psi - \theta - \phi$), la matrice de rotation s'exprime comme :

$$R = \begin{bmatrix} c\psi c\theta & c\psi s\theta s\phi - s\psi c\phi & c\psi s\theta c\phi - s\psi s\phi \\ s\psi c\theta & s\psi s\theta s\phi + c\psi c\phi & s\psi s\theta c\phi - c\psi s\phi \\ -s\theta & c\theta s\phi & c\theta c\phi \end{bmatrix}$$

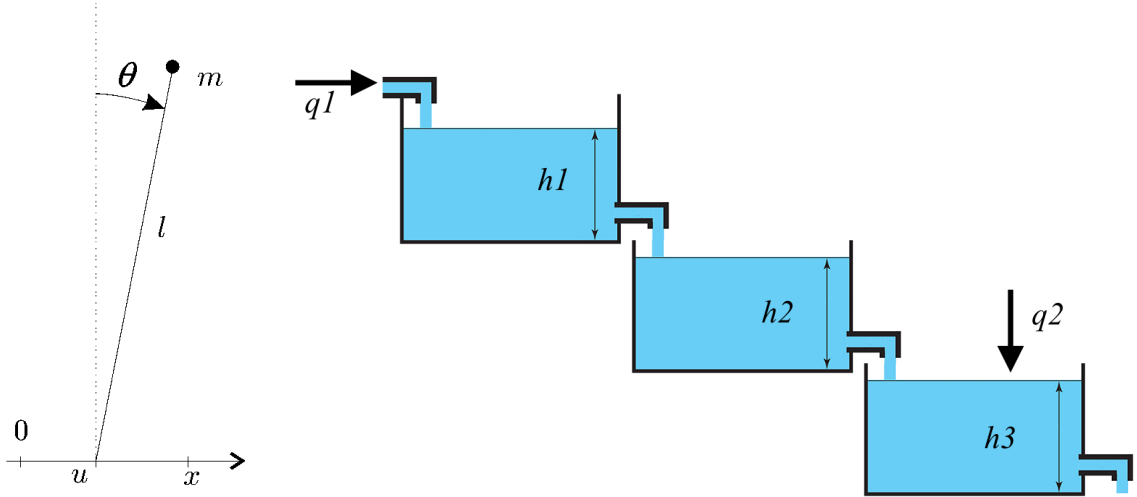
- $\Omega = [\omega_x, \omega_y, \omega_z]^T$ est le vecteur des vitesses angulaires en repère drone. La matrice d'inertie est supposée diagonale $J = \text{diag}(J_1, J_2, J_3)$. Les couples Γ sont fonctions des vitesses de rotation respectives des moteurs.

4.4 Le pendule (inverse)

L'équation du pendule inverse est non linéaire. On peut toutefois la linéariser pour obtenir un système LTI, sous hypothèse de petites oscillations. L'équation de la physique linéarisée (sur l'axe x) est : $\ddot{x} = \frac{g}{l}(x - u)$, et l'on mesure $y = x$. Le choix des variables d'état est alors : $x_1 = x$ et $x_2 = \dot{x}$, ce qui donne la représentation suivante :

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{g}{l} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\frac{g}{l} \end{bmatrix} u$$

$$y = [1, 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



4.5 3 cuves

On considère un système composé de 3 cuves en cascade. L'objectif est d'estimer et réguler les trois hauteurs d'eau h_1, h_2, h_3 à l'aide des deux débits d'entrée q_1, q_2 .

Toutes les cuves sont identiques, de section circulaire S et on fait l'hypothèse que le débit sortant est proportionnel à la hauteur d'eau avec une constante $\alpha > 0$. Le modèle dynamique du système est alors :

$$\begin{cases} S\dot{h}_1 = q_1 - \alpha h_1 \\ S\dot{h}_2 = \alpha h_1 - \alpha h_2 \\ S\dot{h}_3 = q_2 + \alpha h_2 - \alpha h_3 \end{cases}$$

5 Eléments de mathématiques pour l'ingénieur

5.1 Algèbre matricielle

5.1.1 Valeurs propres

Soit une matrice carrée \mathbf{A} , de dimension $n \times n$. λ est valeur propre de \mathbf{A} s'il existe un vecteur \mathbf{u} de \mathbb{R}^n non nul tel que $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$; \mathbf{u} est alors appelé vecteur propre de \mathbf{A} associé à la valeur propre λ .

Les valeurs propres de \mathbf{A} sont les solutions en λ de l'équation (polynôme caractéristique)

$$\det(\mathbf{A} - \lambda\mathbf{I}_n) = 0$$

et les vecteurs propres correspondants sont les éléments du noyau de $\mathbf{A} - \lambda\mathbf{I}_n$.

5.1.2 Calcul du déterminant d'une matrice

Il faut donc savoir calculer le déterminant d'une matrice pour être en mesure de calculer ses valeurs propres.

Déterminant d'une matrice de dimension 2 Le déterminant de

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

est donné par

$$\det \mathbf{A} = ad - bc$$

Déterminant d'une matrice de dimension 3 Il faut imaginer la configuration suivante :

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

Le déterminant de la matrice

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

se calcule alors en développant par rapport à n'importe quelle ligne ou n'importe quelle colonne, c'est-à-dire en appliquant à chaque coefficient le signe correspondant à la matrice ci-dessus, et en multipliant celui-ci par le déterminant de la matrice de dimension 2 obtenue en rayant la ligne et la colonne contenant le terme considéré.

Par exemple, le calcul du déterminant de \mathbf{A} en développant par rapport à sa première ligne est :

$$\det \mathbf{A} = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$

$$\det \mathbf{A} = aei - afh - bdi + bfg + cdh - ceg$$

Déterminant d'une matrice de dimension quelconque La procédure appliquée en dimension 3 se généralise facilement.

5.1.3 Inversion de matrice

L'inverse \mathbf{A}^{-1} d'une matrice \mathbf{A} est telle que $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n$.

Inverse d'une matrice de dimension 2 Soit $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. L'inverse \mathbf{A}^{-1} de \mathbf{A} est donnée par

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Inverse d'une matrice de dimension 3 Soit $\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, son inverse est donnée par

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{bmatrix}$$

Inverse d'une matrice diagonale de dimension quelconque L'inverse d'une matrice diagonale s'obtient en prenant l'inverse de ses termes diagonaux.

Soit $\mathbf{A} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, son inverse est donnée par $\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{a} & 0 & 0 \\ 0 & \frac{1}{b} & 0 \\ 0 & 0 & \frac{1}{c} \end{bmatrix}$

5.2 Probabilités

Le bruit de mesure est fréquemment modélisé comme *blanc Gaussien*, ce qui signifie qu'il suit une *loi normale* de moyenne nulle et de variance donnée.

Une loi de probabilité permet de représenter le comportement d'une variable aléatoire. Autrement dit, de représenter "combien de chance" il y a d'obtenir la valeur de la variable aléatoire dans un intervalle donné. Par exemple, la *loi normale*, dite encore *Gaussienne*, a l'allure suivante

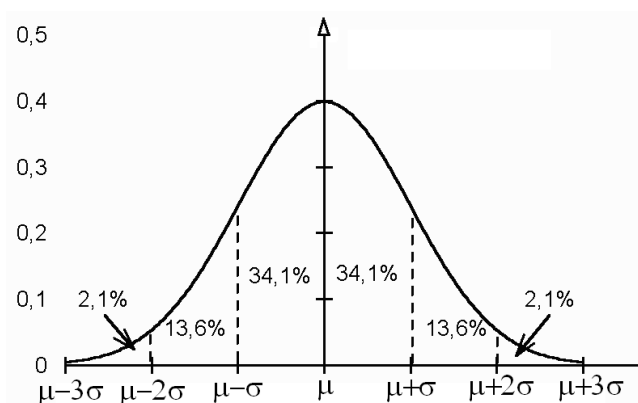


FIGURE 1 : Loi normale, Gaussienne, courbe en cloche

Cette figure signifie que la valeur la plus *probable* pour la valeur aléatoire est sa moyenne μ . L'intégrale entre deux valeurs donne le pourcentage de chance (la *probabilité*) que la variable aléatoire prenne effectivement une valeur dans un intervalle donné. Par exemple ici, il y a 68.2% de probabilité pour que la variable aléatoire se situe entre $\mu - \sigma$ et $\mu + \sigma$ (σ est l'écart type). Il y a de même 95% de chance d'être entre $\mu - 2\sigma$ et $\mu + 2\sigma$ et 99.7% de chance d'être entre $\mu - 3\sigma$ et $\mu + 3\sigma$.

Espérance Le concept de moyenne est bien connu. Pour une variable aléatoire réelle X , elle correspond à l'*espérance*

$$\mu = \mathbb{E}[X]$$

Variance La variance, quant à elle, est une caractérisation de la *dispersion*, qui se calcule comme suit

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]$$

L'écart type σ est égal à la racine carrée de la variance, $\sigma = \sqrt{\text{Var}(X)}$.

Le bruit, phénomène de moyenne nulle, est donc fréquemment représenté suivant une loi normale. Cela signifie que la moyenne 0 est la valeur la plus probable, mais qu'une dispersion aléatoire (de support théorique infini) existe autour de cette valeur moyenne nulle. Outre ces considérations physiques, ce choix est notamment dû au fait que la loi normale ne dépend que de deux paramètres.

Covariance Soit deux variables aléatoires X et Y . La covariance permet d'évaluer le sens de variation de deux variables et de qualifier ainsi leur degré d'indépendance l'une par rapport à l'autre (leur covariance est nulle si elles sont indépendantes). La covariance mesure la variation simultanée de X et Y . Plus elle devient positive, plus cela signifie que les valeurs prises par X et Y sont dispersées de leur moyenne dans le même sens.

$$\text{Cov}(X, Y) = \sigma_{XY} = \mathbb{E}[(X - \mathbb{E}(X))](Y - \mathbb{E}(Y))$$

Notez que $\sigma_{XX} = \text{Var}(X)$.

5.3 Optimisation

Le problème de l'optimisation (sans contrainte) consiste à trouver la valeur $\hat{\theta}$ des paramètres $\theta \in \mathbb{X}$ minimisant (ou maximisant) une fonction scalaire réelle $J(\theta)$. Cela se formule comme

$$\hat{\theta} = \arg \min_{\theta \in \mathbb{X}} J(\theta).$$

Les critères d'optimalité pour un minimum local au point $\theta = \theta_{\text{opt}}$ sont donnés par :

1. Le gradient du critère est nul, cela signale un extremum (minimum ou maximum) local,

$$\left[\frac{\partial J(\theta)}{\partial \theta} \right]_{\theta=\theta_{\text{opt}}} = \mathbf{0}$$

2. Le Hessien du critère est positif au point $\theta = \theta_{\text{opt}}$, cela signale un minimum local,

$$\left[\frac{\partial^2 J(\theta)}{\partial \theta \partial \theta^T} \right]_{\theta=\theta_{\text{opt}}} > \mathbf{0}$$

La réunion de ces deux critères signale un minimum local. À noter que pour maximiser une fonction, il suffit de minimiser la fonction opposée.

Un minimum global peut être obtenu si le critère est *convexe*, car le seul minimum local possible est unique et donc *a fortiori* global. Dans le cas d'un critère non convexe, cette technique permet simplement de caractériser les minima locaux et il faut recourir à des algorithmes plus complexes qui ne seront pas détaillés ici pour trouver le minimum global.

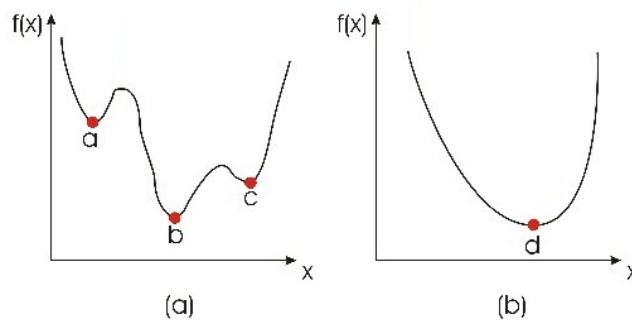


FIGURE 2 : Minimum local et global. (a) Non convexe, minima locaux. (b) Convexe, minimum global