



Mathématiques appliquées – Analyse numérique L2 Géosciences fondamentales

Stéphane Jacquemoud (10-mai-06)

I. Rappels sur les matrices

I.1. Définitions

- On appelle matrice A ou $A_{p \times q}$ un **tableau rectangulaire** constitué de $p \times q$ éléments génériques a_{ij} , $i \in \{1, \dots, p\}$ et $j \in \{1, \dots, q\}$. Les éléments a_{ij} sont des nombres réels ou complexes :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{pmatrix}$$

Le nombre de lignes p et le nombre de colonnes q sont les **dimensions** de la matrice.

- Une matrice qui ne comporte qu'une seule colonne est une **matrice unicolonne** de dimension $p \times 1$ ou **vecteur colonne**.
- Une matrice qui ne comporte qu'une seule ligne est une **matrice uniligne** de dimension $1 \times q$ ou **vecteur ligne**.
- La **matrice nulle** de dimension $p \times q$ est une matrice dont tous les éléments a_{ij} sont nuls.
- La **matrice unité** de dimension $p \times p$ est une matrice dont tous les éléments a_{ij} sont égaux à 1.
- La **transposée** A^T de la matrice A de dimension $p \times q$ est une matrice de dimension $q \times p$ obtenue en échangeant les lignes et les colonnes de A .

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{p1} \\ a_{12} & a_{22} & \cdots & a_{p2} \\ \vdots & \vdots & & \vdots \\ a_{1q} & a_{2q} & \cdots & a_{pq} \end{pmatrix}$$

- Une **matrice carrée** est une matrice de dimension $p \times p$ dont le nombre de lignes est égal au nombre de colonnes.

Cas particulier des matrices carrées

- Une **matrice triangulaire** est une matrice dont tous les éléments situés d'un côté de la diagonale principale sont nuls : **matrice triangulaire inférieure** ($a_{ij} = 0, \forall j > i$) et **matrice triangulaire supérieure** ($a_{ij} = 0, \forall i > j$).
- On parle de **matrice symétrique** lorsque tous les éléments occupant des positions symétriques par rapport à la diagonale principale sont égaux. Pour qu'une matrice carrée soit symétrique, il faut et il suffit qu'elle soit égale à sa transposée.
- Une **matrice diagonale** est une matrice dont tous les éléments situés en dehors la diagonale sont nuls.
- La **matrice identité** I_p est une matrice diagonale dont tous les éléments de la diagonale principale valent 1.

I.2. Opérations sur les matrices

Soit B une autre matrice constituée de $p \times q$ éléments génériques b_{ij} :

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pq} \end{pmatrix}$$

- **Somme de deux matrices** : on appelle somme de deux matrices A et B de dimension $p \times q$ la matrice C de même dimension dont les éléments sont donnés par $c_{ij} = a_{ij} + b_{ij}$.
- **Produit d'une matrice par un scalaire** : on appelle produit de A par le scalaire λ la matrice C de même dimension dont les éléments sont donnés par $c_{ij} = \lambda a_{ij}$.
- **Produit matriciel** : on appelle produit d'une matrice A de dimension $p \times q$ par une matrice B de dimension $q \times r$ la matrice C de dimension $p \times r$ dont les éléments sont donnés par $c_{ij} = \sum_{k=1}^q a_{ik} \times b_{kj}$ avec $i \in \{1, \dots, p\}$ et $j \in \{1, \dots, r\}$. Le produit matriciel n'est pas commutatif : $AB \neq BA$. Il est associatif : $(AB)C = A(BC)$. Enfin il est distributif à gauche et à droite par rapport à l'addition : $A(B+C) = AB+AC$ et $(B+C)A = BA+CA$.

Ex : Soit $A = \begin{pmatrix} 1 & -1 & 0 & -3 \\ 3 & -2 & 1 & 0 \end{pmatrix}$. Calculer $A^T A$ et AA^T .

$$A^T A = \begin{pmatrix} 1 & 3 \\ -1 & -2 \\ 0 & 1 \\ -3 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & -1 & 0 & -3 \\ 3 & -2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 10 & -7 & 3 & -3 \\ -7 & 5 & -2 & 3 \\ 3 & -2 & 1 & 0 \\ -3 & 3 & 0 & 9 \end{pmatrix} \text{ et } AA^T = \begin{pmatrix} 1 & -1 & 0 & -3 \\ 3 & -2 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & 3 \\ -1 & -2 \\ 0 & 1 \\ -3 & 0 \end{pmatrix} = \begin{pmatrix} 11 & 5 \\ 5 & 14 \end{pmatrix}$$

Ex : Soit $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Calculer A^n .

$$A^2 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, A^3 = \begin{pmatrix} 1 & 3 & 3 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, A^4 = \begin{pmatrix} 1 & 4 & 6 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}. \text{ On suppose que } A^n = \begin{pmatrix} 1 & n & \frac{n(n-1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} \text{ et on montre}$$

par récurrence que cette relation est valable au rang $n+1$.

I.3. Déterminant d'une matrice carrée

Soit A une matrice carrée d'ordre 2 d'éléments a_{ij} : $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

Le déterminant de A s'écrit : $\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$

Pour calculer le déterminant de la matrice $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ carrée d'ordre 3, on peut développer suivant la première

$$\text{colonne : } \det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{2+1} a_{12} \begin{vmatrix} a_{21} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + (-1)^{3+1} a_{13} \begin{vmatrix} a_{21} & a_{12} \\ a_{32} & a_{23} \end{vmatrix}$$

Considérons maintenant une matrice A carrée d'ordre p : on appelle **mineur** de A le déterminant de la sous-matrice extraite de A à laquelle on a ôté la $j^{\text{ème}}$ colonne et la $i^{\text{ème}}$ ligne. On le note généralement Δ_{ij} . On appelle **matrice des cofacteurs** ou **comatrice** notée $com(A)$ la matrice de terme général $\gamma_{ij} = (-1)^{i+j} \Delta_{ij}$. Pour une matrice carrée d'ordre 3 elle s'écrit donc :

$$\text{com}(A) = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix} = \begin{pmatrix} (-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & (-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & (-1)^{1+3} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ (-1)^{2+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & (-1)^{2+2} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & (-1)^{2+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ (-1)^{3+1} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} & (-1)^{3+2} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & (-1)^{3+3} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix}$$

Le déterminant de A s'écrit alors : $\det(A) = \sum_{i=1}^p a_{ij} \gamma_{ij} = \sum_{i=1}^p (-1)^{i+j} a_{ij} \Delta_{ij}$ suivant la j^{ème} colonne et

$$\det(A) = \sum_{j=1}^p a_{ij} \gamma_{ij} = \sum_{j=1}^p (-1)^{i+j} a_{ij} \Delta_{ij} \text{ suivant la i^{ème} ligne.}$$

On peut montrer que $\det(A) = \det(A^T)$ et que $\det(AB) = \det(BA) = \det(A) \times \det(B)$.

Ex : Calculer le déterminant de $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 4 & 2 & 5 \end{pmatrix}$.

$$\det(A) = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 4 & 2 & 5 \end{vmatrix} = 1 \times \begin{vmatrix} 3 & 4 \\ 2 & 5 \end{vmatrix} - 2 \times \begin{vmatrix} 2 & 3 \\ 2 & 5 \end{vmatrix} + 4 \times \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} = -5$$

I.4. Inverse d'une matrice carrée

Une matrice carrée A d'ordre p est dite **inversible** s'il existe une matrice carrée A^{-1} de même ordre telle que $AA^{-1} = A^{-1}A = I_p$ avec I_p la matrice identité. A^{-1} est la **matrice inverse** de A. Une condition pour qu'une matrice carrée soit inversible est que $\det(A) \neq 0$ et que le rang de A soit égal à p. Parmi les méthodes permettant de calculer la matrice inverse, celle utilisant la **comatrice** est très générale. On peut montrer que la transposée de la comatrice de A est égale à la comatrice de la transposée de A : $(\text{com}(A))^T = \text{com}(A^T)$. Si A est inversible, alors $A^{-1} = \frac{\text{com}(A^T)}{\det(A)}$.

Ex : Inverser $A = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}$.

$$A^T = \begin{pmatrix} 4 & 2 \\ 5 & 3 \end{pmatrix} \Rightarrow \text{com}(A^T) = \begin{pmatrix} 3 & -5 \\ -2 & 4 \end{pmatrix} \text{ et } \det(A) = \begin{vmatrix} 4 & 5 \\ 2 & 3 \end{vmatrix} = 12 - 10 = 2 \text{ donc } A^{-1} = \frac{\text{com}(A^T)}{\det(A)} = \begin{pmatrix} 3 & -5 \\ -2 & 4 \end{pmatrix}$$

$$\text{On vérifie bien que } A \times A^{-1} = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix} \times \begin{pmatrix} 3 & -5 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

I.5. Valeurs propres et vecteurs propres

Les valeurs propres et les vecteurs propres jouent un rôle important dans le calcul matriciel pour diverses applications : résolution de systèmes d'équations différentielles, analyse en composantes principales.

- Considérons deux matrices A et B de dimension $p \times q$: A et B sont dites **équivalentes** s'il existe deux matrices carrées P et Q d'ordres respectifs p et q telles que $A = PBQ$.
- Considérons maintenant deux matrices carrées A et B de même ordre p. A et B sont dites **semblables** s'il existe une matrice carrée P d'ordre p telle que $A = PBP^{-1}$.
- Par définition on dit que la matrice carrée A est **diagonalisable** s'il existe une matrice diagonale B semblable à A. La diagonale de B est formée par les **valeurs propres** de la matrice A, et P par ses **vecteurs propres**. La diagonalisation d'une matrice permet de calculer sa puissance n^{ième}.

- Les valeurs propres λ de la matrice carrée A d'ordre p sont les solutions de l'équation $\phi(\lambda) = \det(A - \lambda I_p) = 0$ appelée **équation caractéristique**. $\phi(\lambda)$ est un polynôme de degré p dont nous verrons plus tard qu'il admet p racines réelles ou complexes. L'ensemble des valeurs propres de A est dit **spectre** de A.
- A chacune valeur propre λ_α pour $\alpha \in \{1, 2, \dots, p\}$ on peut associer un vecteur propre \vec{v}_α solution de l'équation $A\vec{v}_\alpha = \lambda_\alpha \vec{v}_\alpha \Leftrightarrow (A - \lambda_\alpha I_p)\vec{v}_\alpha = \vec{0}_{p \times 1}$ où $\vec{0}_{p \times 1}$ est le vecteur colonne nul de dimension $p \times 1$.

Ex : Soit la matrice carrée $A = \begin{pmatrix} 5 & 1 & -1 \\ 2 & 4 & -2 \\ 1 & -1 & 3 \end{pmatrix}$. Diagonaliser A puis calculer A^n .

$$\det(A - \lambda I) = \begin{vmatrix} 5-\lambda & 1 & -1 \\ 2 & 4-\lambda & -2 \\ 1 & -1 & 3-\lambda \end{vmatrix} = (5-\lambda) \times \begin{vmatrix} 4-\lambda & -2 \\ -1 & 3-\lambda \end{vmatrix} - 2 \times \begin{vmatrix} 1 & -1 \\ -1 & 3-\lambda \end{vmatrix} + 1 \times \begin{vmatrix} 1 & -1 \\ 4-\lambda & -2 \end{vmatrix}$$

$$= -\lambda^3 + 12\lambda^2 - 44\lambda + 48 = -(\lambda - 4)(\lambda - 2)(\lambda - 6)$$

On obtient trois valeurs propres distinctes classées par ordre croissant : $\lambda_1 = 2$, $\lambda_2 = 4$ et $\lambda_3 = 6$.

$$\text{Soit } \lambda_1 = 2 : (A - 2I)\vec{v}_1 = 0 \Rightarrow \begin{cases} 3x + y - z = 0 \\ 2x + 2y - 2z = 0 \\ x - y + z = 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = z \end{cases} \Rightarrow \vec{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$\text{Soit } \lambda_2 = 4 : (A - 4I)\vec{v}_2 = 0 \Rightarrow \begin{cases} x + y - z = 0 \\ 2x - 2z = 0 \\ x - y - z = 0 \end{cases} \Rightarrow \begin{cases} x = z \\ y = 0 \end{cases} \Rightarrow \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$\text{Soit } \lambda_3 = 6 : (A - 6I)\vec{v}_3 = 0 \Rightarrow \begin{cases} -x + y - z = 0 \\ 2x - 2y - 2z = 0 \\ x - y - 3z = 0 \end{cases} \Rightarrow \begin{cases} x = y \\ z = 0 \end{cases} \Rightarrow \vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\text{Donc } B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix} \text{ et } P = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \Rightarrow P^{-1} = \frac{1}{2} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

On écrit que $A = PBP^{-1} \Rightarrow A^2 = (PBP^{-1})(PBP^{-1}) = PB^2P^{-1}$

On montre de même que $A^3 = PB^3P^{-1}$ et on généralise cette formule au rang n par récurrence : $A^n = PB^nP^{-1}$

$$A^n = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2^n & 0 & 0 \\ 0 & 4^n & 0 \\ 0 & 0 & 6^n \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 4^n + 6^n & -4^n + 6^n & 4^n - 6^n \\ -2^n + 6^n & 2^n + 6^n & 2^n - 6^n \\ -2^n + 4^n & 2^n - 4^n & 2^n + 4^n \end{pmatrix}$$

Remarque : la méthode employée dans l'exemple ci-dessus ne convient pas pour des matrices carrées d'ordre supérieur à trois ou quatre. Il existe deux grandes classes de méthodes alternatives permettant de déterminer les valeurs propres : les **méthodes directes** (méthode des rotations de Jacobi, réduction de Householder, algorithme QL, etc.) et les **méthodes itératives** (algorithme de la puissance itérée, méthode de Lanczos, etc.)

Considérons maintenant une matrice carrée A **réelle** et **symétrique** : on admettra que ses valeurs propres sont réelles et que les vecteurs propres correspondant à des valeurs propres distinctes sont orthogonaux. La matrice carrée P est dite **orthogonale**, c'est-à-dire $P^T P = P P^T = I \Rightarrow P^{-1} = P^T$. Donc $A = PBP^{-1} \Rightarrow B = P^{-1}AP = P^T AP$. Cette relation permet de déterminer les valeurs propres de A sans résoudre l'équation caractéristique. Considérons la matrice carrée

$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. On peut montrer facilement que R est orthogonale quelle que soit la valeur de θ .

$$R^T R = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \times \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} \cos^2 \theta + \sin^2 \theta & -\cos \theta \sin \theta + \cos \theta \sin \theta \\ -\cos \theta \sin \theta + \cos \theta \sin \theta & \cos^2 \theta + \sin^2 \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$R R^T = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \times \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} \cos^2 \theta + \sin^2 \theta & \cos \theta \sin \theta - \cos \theta \sin \theta \\ \cos \theta \sin \theta - \cos \theta \sin \theta & \cos^2 \theta + \sin^2 \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Le produit de R par un vecteur colonne est une **rotation**. Si ce vecteur est unitaire, cette opération revient à déplacer un point le long d'un cercle. Calculons maintenant le produit matriciel $R^T A R$:

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a_{11} \cos \theta + a_{12} \sin \theta & -a_{11} \sin \theta + a_{12} \cos \theta \\ a_{21} \cos \theta + a_{22} \sin \theta & -a_{21} \sin \theta + a_{22} \cos \theta \end{pmatrix}$$

$$= \begin{pmatrix} (a_{11} \cos \theta + a_{12} \sin \theta) \cos \theta + (a_{21} \cos \theta + a_{22} \sin \theta) \sin \theta & (-a_{11} \sin \theta + a_{12} \cos \theta) \cos \theta + (-a_{21} \sin \theta + a_{22} \cos \theta) \sin \theta \\ -(a_{11} \cos \theta + a_{12} \sin \theta) \sin \theta + (a_{21} \cos \theta + a_{22} \sin \theta) \cos \theta & (-a_{11} \sin \theta + a_{12} \cos \theta) \sin \theta + (-a_{21} \sin \theta + a_{22} \cos \theta) \cos \theta \end{pmatrix}$$

On cherche la valeur de θ telle que cette matrice est diagonale. Comme $a_{21} = a_{12}$ cela revient à poser que :

$$-(a_{11} \cos \theta + a_{12} \sin \theta) \sin \theta + (a_{21} \cos \theta + a_{22} \sin \theta) \cos \theta = 0$$

$$\Rightarrow (a_{22} - a_{11}) \sin \theta \cos \theta - a_{12} \sin^2 \theta + a_{21} \cos^2 \theta = 0$$

$$\Rightarrow (a_{22} - a_{11}) \sin \theta \cos \theta + a_{12} (\cos^2 \theta - \sin^2 \theta) = 0$$

$$\Rightarrow (a_{22} - a_{11}) \frac{\sin 2\theta}{2} + a_{12} \cos 2\theta = 0$$

On considère deux cas :

$$\begin{cases} \text{si } a_{22} \neq a_{11} \Rightarrow \tan 2\theta = -\frac{2a_{12}}{a_{22} - a_{11}} \Rightarrow \theta = \frac{1}{2} \arctan\left(-\frac{2a_{12}}{a_{22} - a_{11}}\right) = -\frac{1}{2} \arctan\left(\frac{2a_{12}}{a_{22} - a_{11}}\right) \\ \text{si } a_{22} = a_{11} \Rightarrow \cos 2\theta = 0 \Rightarrow \theta = \frac{\pi}{4} \end{cases}$$

Cette opération consistant à transformer une matrice symétrique en une matrice diagonale est appelée **rotation de Jacobi**. On peut la généraliser à une matrice carrée réelle et symétrique d'ordre n : elle permet alors d'annuler un à un tous les éléments de la matrice situés en dehors de la diagonale principale. Une rotation n'affectant pas les valeurs propres, et une succession de rotations équivalant à une seule rotation, la diagonale finale contient les valeurs propres !

Ex : Calculer les valeurs propres de la matrice carrée $A = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$ par deux méthodes différentes.

$$\det(A - \lambda I) = \begin{vmatrix} 5 - \lambda & 2 \\ 2 & 5 - \lambda \end{vmatrix} = (5 - \lambda)^2 - 4 = (5 - \lambda - 2)(5 - \lambda + 2) = (3 - \lambda)(7 - \lambda)$$

$$a_{22} = a_{11} \Rightarrow \theta = \frac{\pi}{4} \text{ donc } B = P^T A P = \begin{pmatrix} \cos \frac{\pi}{4} & \sin \frac{\pi}{4} \\ -\sin \frac{\pi}{4} & \cos \frac{\pi}{4} \end{pmatrix} \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} \cos \frac{\pi}{4} & -\sin \frac{\pi}{4} \\ \sin \frac{\pi}{4} & \cos \frac{\pi}{4} \end{pmatrix} = \begin{pmatrix} 7 & 0 \\ 0 & 3 \end{pmatrix}$$

Donc 3 et 7 sont les deux valeurs propres de A.

1.6. Application des matrices à la résolution d'un système linéaire de q équations à p inconnues x_1, \dots, x_p ($q \geq p$)

Considérons le système linéaire

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1j}x_j + \dots + a_{1p}x_p = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2j}x_j + \dots + a_{2p}x_p = b_2 \\ \dots \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ij}x_j + \dots + a_{ip}x_p = b_i \\ \dots \\ a_{q1}x_1 + a_{q2}x_2 + \dots + a_{qj}x_j + \dots + a_{qp}x_p = b_q \end{cases}$$

Où les valeurs a_{ij} et b_i sont **connues** et où les x_i sont **inconnus**. Pour résoudre ce problème, on écrit le système sous forme matricielle :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{q1} & a_{q2} & \dots & a_{qj} & \dots & a_{qp} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_q \end{pmatrix} \Leftrightarrow AX = B$$

- Si $p = q$ alors A est une matrice carrée. Il existe une infinité de solutions si $\det(A) = 0$ et une solution unique si $\det(A) \neq 0$: dans ce cas A^{-1} existe et $X = A^{-1}B$.

- Si $p < q$ alors il n'y a pas de solution unique à ce système car le nombre d'équations q est supérieur au nombre d'inconnues p . On va donc déterminer la solution qui satisfait au mieux l'ensemble des équations. La matrice A qui a pour dimension $p \times q$ n'est pas carrée et donc pas inversible. On se ramène à une matrice carrée en multipliant A par sa matrice transposée A^T :

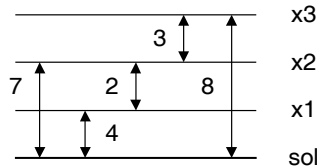
$$AX = B \Rightarrow A^T (AX) = A^T B \Rightarrow (A^T A) X = A^T B$$

Le produit d'une matrice de dimension $p \times q$ par sa transposée de dimension $q \times p$ est une matrice carrée d'ordre p dont on peut calculer l'inverse. On obtient alors :

$$(A^T A) X = A^T B \Rightarrow (A^T A)^{-1} (A^T A) X = (A^T A)^{-1} A^T B \Rightarrow IX = (A^T A)^{-1} A^T B \Rightarrow X = (A^T A)^{-1} A^T B$$

Cette transformation s'appelle l'**élimination gaussienne**.

Ex : Cinq mesures permettent de déterminer les hauteurs x_1 , x_2 et x_3 de trois niveaux. Calculer ces hauteurs par moindres carrés.



On pose le système d'équations linéaires suivant :

$$\begin{cases} x_1 = 4 \\ x_2 = 7 \\ x_3 = 8 \\ x_2 - x_1 = 2 \\ x_3 - x_2 = 3 \end{cases} \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 8 \\ 2 \\ 3 \end{pmatrix} \Rightarrow AX = B$$

On sait que la solution X s'écrit : $X = (A^T A)^{-1} A^T B$

$$A^T A = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

$$\det(A^T A) = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 8 \text{ et } \text{com}\left((A^T A)^T\right) = \text{com}\left(\begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{pmatrix}\right) = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 5 \end{pmatrix}$$

$$\text{Donc } X = \frac{1}{8} \begin{pmatrix} 5 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 5 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 4 \\ 7 \\ 8 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 5 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 5 \end{pmatrix} \times \begin{pmatrix} 2 \\ 6 \\ 11 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 33 \\ 50 \\ 69 \end{pmatrix} = \begin{pmatrix} 4.125 \\ 6.25 \\ 8.625 \end{pmatrix}$$

I.7. Application des matrices à la résolution d'un système non-linéaire de q équations à p inconnues $x_1 \dots x_p$ ($q > p$)

$$\text{Considérons le système non-linéaire } \begin{cases} f_1(x_1, x_2, \dots, x_j, \dots, x_p) = b_1 \\ f_2(x_1, x_2, \dots, x_j, \dots, x_p) = b_2 \\ \dots \\ f_i(x_1, x_2, \dots, x_j, \dots, x_p) = b_i \\ \dots \\ f_q(x_1, x_2, \dots, x_j, \dots, x_p) = b_q \end{cases}, \text{ par exemple } \begin{cases} 2x_1 x_2 + x_3^2 = 8 \\ x_1 + x_1 x_2 = 6 \\ 2x_1 x_3 = 5 \\ 4x_1 + 2x_2 = 9 \end{cases}$$

Ce système n'est pas linéaire si bien qu'on ne peut pas extraire le vecteur des $x_1, x_2, \dots, x_j, \dots, x_p$. La méthode de résolution consiste à développer les fonctions $f_1, f_2, \dots, f_i, \dots, f_q$ en **série de Taylor** au premier ordre.

Rappels sur les développements en série de Taylor

Soit $f(x)$ une fonction de la variable x . Le développement de Taylor de f à l'ordre n en a s'écrit :

$$f(x) = f(a) + \frac{\partial f}{\partial x}(a) \frac{(x-a)}{1!} + \frac{\partial^2 f}{\partial x^2}(a) \frac{(x-a)^2}{2!} + \dots + \frac{\partial^n f}{\partial x^n}(a) \frac{(x-a)^n}{n!} = f(a) + \sum_{k=1}^n \frac{\partial^k f}{\partial x^k}(a) \frac{(x-a)^k}{k!}$$

Soit $f(x, y)$ une fonction de deux variables x et y . Le développement de Taylor de f à l'ordre n en a et b s'écrit :

$$f(x, y) = f(a, b) + (x-a) \frac{\partial f}{\partial x}(a, b) + (y-b) \frac{\partial f}{\partial y}(a, b) + \frac{1}{2!} \left\{ (x-a)^2 \frac{\partial^2 f}{\partial x^2}(a, b) + 2(x-a)(y-b) \frac{\partial^2 f}{\partial x \partial y}(a, b) + (y-b)^2 \frac{\partial^2 f}{\partial y^2}(a, b) \right\} + \dots$$

où $\frac{\partial f}{\partial x}(a, b) = f_x(a, b)$, $\frac{\partial f}{\partial y}(a, b) = f_y(a, b)$ désignent les dérivées partielles par rapport à x et y calculées en $x = a$ et $y = b$. On peut généraliser cette expression pour une fonction de plusieurs variables $f(x_1, x_2, \dots, x_j, \dots, x_p)$.

Linéarisation du système

En développant en série de Taylor au premier ordre chaque fonction f de notre système non-linéaire, on obtient :

$$\begin{cases} f_1(a_1, a_2, \dots, a_p) + \sum_{j=1}^p (x_j - a_j) \frac{\partial f_1}{\partial x_j}(a_1, a_2, \dots, a_p) = b_1 \\ f_2(a_1, a_2, \dots, a_p) + \sum_{j=1}^p (x_j - a_j) \frac{\partial f_2}{\partial x_j}(a_1, a_2, \dots, a_p) = b_2 \\ \dots \\ f_q(a_1, a_2, \dots, a_p) + \sum_{j=1}^p (x_j - a_j) \frac{\partial f_q}{\partial x_j}(a_1, a_2, \dots, a_p) = b_q \end{cases}$$

qui peut s'écrire sous forme matricielle :

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a_1, a_2, \dots, a_p) & \frac{\partial f_1}{\partial x_2}(a_1, a_2, \dots, a_p) & \dots & \frac{\partial f_1}{\partial x_p}(a_1, a_2, \dots, a_p) \\ \frac{\partial f_2}{\partial x_1}(a_1, a_2, \dots, a_p) & \frac{\partial f_2}{\partial x_2}(a_1, a_2, \dots, a_p) & \dots & \frac{\partial f_2}{\partial x_p}(a_1, a_2, \dots, a_p) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_q}{\partial x_1}(a_1, a_2, \dots, a_p) & \frac{\partial f_q}{\partial x_2}(a_1, a_2, \dots, a_p) & \dots & \frac{\partial f_q}{\partial x_p}(a_1, a_2, \dots, a_p) \end{pmatrix} \times \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_p - a_p \end{pmatrix} = \begin{pmatrix} b_1 - f_1(a_1, a_2, \dots, a_p) \\ b_2 - f_2(a_1, a_2, \dots, a_p) \\ \vdots \\ b_q - f_q(a_1, a_2, \dots, a_p) \end{pmatrix}$$
$$\Rightarrow A \times \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_p - a_p \end{pmatrix} = B \Rightarrow \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_p - a_p \end{pmatrix} = (A^T A)^{-1} A^T B \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = (A^T A)^{-1} A^T B + \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}$$

Ex : Résoudre le système suivant en partant du point $(x_0, y_0) = (3.1, 0.1)$ qui semble proche de la solution

$$\begin{cases} (x-2)^2 - \frac{1}{3}y^2 = 1 \\ (y-2)^2 - \frac{1}{3}x^2 = 1 \\ \frac{1}{8}(x-y)^2 + \frac{1}{8}(x+y-4)^2 = 1 \end{cases} \Leftrightarrow \begin{cases} f_1(x, y) = x^2 - 4x - \frac{1}{3}y^2 + 4 = 1 \\ f_2(x, y) = y^2 - 4y - \frac{1}{3}x^2 + 4 = 1 \\ f_3(x, y) = \frac{1}{4}x^2 + \frac{1}{4}y^2 - x - y + 2 = 1 \end{cases}$$

On calcule les dérivées partielles :

$$\frac{\partial f_1}{\partial x} = 2x - 4, \quad \frac{\partial f_1}{\partial y} = -\frac{2}{3}y, \quad \frac{\partial f_2}{\partial x} = -\frac{2}{3}x, \quad \frac{\partial f_2}{\partial y} = 2y - 4, \quad \frac{\partial f_3}{\partial x} = \frac{1}{2}x - 1 \quad \text{et} \quad \frac{\partial f_3}{\partial y} = \frac{1}{2}y - 1$$

Au point $(x_0, y_0) = (3.1, 0.1)$ la matrice A des dérivées s'écrit donc :

$$A = \begin{pmatrix} 2.200 & -0.067 \\ -2.067 & -3.800 \\ 0.550 & -0.950 \end{pmatrix}, \quad A^T = \begin{pmatrix} 2.200 & -2.067 & 0.550 \\ -0.067 & -3.800 & -0.950 \end{pmatrix}, \quad A^T A = \begin{pmatrix} 9.414 & 7.184 \\ 7.184 & 15.347 \end{pmatrix}$$

$$\text{com}(A^T A) = \begin{pmatrix} 15.347 & -7.184 \\ -7.184 & 9.414 \end{pmatrix} \quad \text{et} \quad \det(A^T A) = \begin{vmatrix} 9.414 & 7.184 \\ 7.184 & 15.347 \end{vmatrix} = 92.858 \Rightarrow (A^T A)^{-1} = \begin{pmatrix} 0.1653 & -0.0774 \\ -0.0774 & 0.1014 \end{pmatrix}$$

$$\text{Le vecteur } B \text{ s'écrit : } B = \begin{pmatrix} 1 - f_1(x_0, y_0) \\ 1 - f_2(x_0, y_0) \\ 1 - f_3(x_0, y_0) \end{pmatrix} = \begin{pmatrix} -0.2067 \\ 0.5933 \\ -0.2050 \end{pmatrix}$$

La solution au problème est donc :

$$\begin{pmatrix} x \\ y \end{pmatrix} = (A^T A)^{-1} A^T B + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0.1653 & -0.0774 \\ -0.0774 & 0.1014 \end{pmatrix} \times \begin{pmatrix} 2.200 & -2.067 & 0.550 \\ -0.067 & -3.800 & -0.950 \end{pmatrix} \times \begin{pmatrix} -0.2067 \\ 0.5933 \\ -0.2050 \end{pmatrix} + \begin{pmatrix} 3.1 \\ 0.1 \end{pmatrix} = \begin{pmatrix} 2.9619 \\ 0.0313 \end{pmatrix}$$

II. Interpolation polynômiale

Considérons un ensemble de $n+1$ points $M_i(x_i, y_i)$, $i \in \{0, 1, \dots, n\}$, qui peuvent être par exemple des résultats de mesures, tels que $x_0 < x_1 < \dots < x_n$. Lorsque l'on cherche à déterminer la valeur y_* en $x_* \neq x_i$, on parle d'**interpolation** si $x_0 < x_* < x_n$ et d'**extrapolation** si $x_* < x_0$ ou $x_* > x_n$. Nous n'aborderons ici que le problème d'interpolation par des polynômes : lorsque les points sont peu nombreux, on cherche à construire la fonction $f(x)$ passant exactement par tous ces points qu'on appelle **points d'appui**. Par deux points passe une droite (polynôme de degré 1) ; par trois points passe une parabole (polynôme de degré 2) ; en généralisant on peut écrire que par $n+1$ points passe un polynôme $P_n(x)$ de degré n : $P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ est appelé **polynôme d'interpolation**. Il existe de nombreuses méthodes pour déterminer $P_n(x)$ – dont celle consistant à poser les $n+1$ équations à $n+1$ inconnues et à inverser la **matrice de Vandermonde** – mais si l'on utilise les mêmes points d'appui, celles-ci doivent fournir le même polynôme.

II.1. Méthode d'interpolation polynômiale de Lagrange

Il s'agit d'une méthode facile à retenir. Considérons les deux points $M_0(x_0, y_0)$ et $M_1(x_1, y_1)$ et cherchons l'équation de la droite les rejoignant. On écrit une fonction de la forme $P(x) = y_0A_0(x) + y_1A_1(x)$ avec $A_0(x)$ et $A_1(x)$ deux polynômes de degré 1 que l'on détermine grâce aux conditions suivantes :

$$\begin{cases} A_0(x) = 1 \text{ quand } x = x_0 \\ A_0(x) = 0 \text{ quand } x = x_1 \end{cases} \quad \text{et} \quad \begin{cases} A_1(x) = 0 \text{ quand } x = x_0 \\ A_1(x) = 1 \text{ quand } x = x_1 \end{cases}$$

Chaque polynôme a pour valeur 1 à l'abscisse d'un point d'appui et s'annule à l'abscisse de l'autre. La fonction s'écrit

$$\text{donc : } P(x) = y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0}$$

En posant $x_1 - x_0 = \Delta x$ on peut faire apparaître le terme $x - x_1$ dans l'expression de $P(x)$:

$$P(x) = y_0 \frac{x-x_1}{x_0-x_1} + y_1 \frac{x-x_1+x_1-x_0}{x_1-x_0} = y_0 \frac{x-x_1}{(-\Delta x)} + y_1 \frac{x-x_1+\Delta x}{\Delta x} = y_1 + (y_1-y_0) \frac{x-x_1}{\Delta x}$$

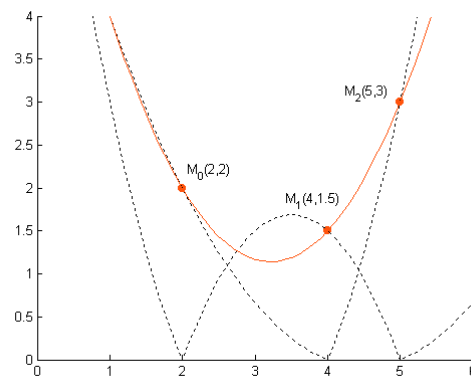
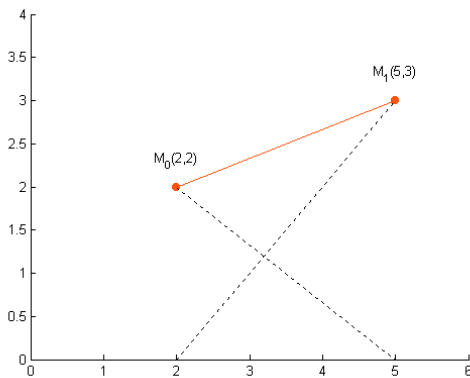
Supposons maintenant que l'on souhaite faire passer une parabole par trois points $M_0(x_0, y_0)$, $M_1(x_1, y_1)$ et $M_2(x_2, y_2)$. On écrit de même $P(x) = y_0 A_0(x) + y_1 A_1(x) + y_2 A_2(x)$ avec $A_0(x)$, $A_1(x)$ et $A_2(x)$ trois polynômes de degré 2 que l'on détermine grâce aux conditions suivantes :

$$\begin{cases} A_0(x) = 1 \text{ quand } x = x_0 \\ A_0(x) = 0 \text{ quand } x = x_1 \text{ ou } x = x_2 \end{cases} \quad \begin{cases} A_1(x) = 1 \text{ quand } x = x_1 \\ A_1(x) = 0 \text{ quand } x = x_0 \text{ ou } x = x_2 \end{cases} \quad \text{et} \quad \begin{cases} A_2(x) = 1 \text{ quand } x = x_2 \\ A_2(x) = 0 \text{ quand } x = x_0 \text{ ou } x = x_1 \end{cases}$$

La fonction s'écrit donc
$$P(x) = y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}$$

Si les abscisses des points sont équidistantes, alors $x_1 - x_0 = x_2 - x_1 = \Delta x$ et $x_2 - x_0 = 2\Delta x$. On transforme $P(x)$ en faisant apparaître le terme $x - x_1$:

$$\begin{aligned} P(x) &= y_0 \frac{(x-x_1)(x-x_1+x_1-x_2)}{(-\Delta x)(-2\Delta x)} + y_1 \frac{(x-x_1+x_1-x_0)(x-x_1+x_1-x_2)}{(\Delta x)(-\Delta x)} + y_2 \frac{(x-x_1+x_1-x_0)(x-x_1)}{(2\Delta x)(\Delta x)} \\ &= y_0 \frac{(x-x_1)(x-x_1-\Delta x)}{(-\Delta x)(-2\Delta x)} + y_1 \frac{(x-x_1+\Delta x)(x-x_1-\Delta x)}{(\Delta x)(-\Delta x)} + y_2 \frac{(x-x_1+\Delta x)(x-x_1)}{(2\Delta x)(\Delta x)} \\ &= y_0 \frac{(x-x_1)^2 - \Delta x(x-x_1)}{2(\Delta x)^2} - y_1 \frac{(x-x_1)^2 - (\Delta x)^2}{(\Delta x)^2} + y_2 \frac{(x-x_1)^2 + \Delta x(x-x_1)}{2(\Delta x)^2} \\ &= y_1 + \frac{y_2 - y_0}{2} \frac{x-x_1}{\Delta x} + \frac{y_2 - 2y_1 + y_0}{2} \frac{(x-x_1)^2}{(\Delta x)^2} \end{aligned}$$



Le développement pour des ordres supérieurs est immédiat. Le $k+1^{\text{ème}}$ terme d'un développement d'ordre n a pour forme

générale :
$$y_k \prod_{j=0, j \neq k}^n \frac{(x-x_j)}{(x_k-x_j)}$$

Remarque : un polynôme de degré trop élevé peut engendrer des problèmes d'oscillation ne correspondant pas à la fonction que l'on souhaite approximer. On choisit alors de construire des approximations par morceaux de faible degré, ces morceaux étant indépendants les uns des autres ou liés par des conditions de continuité de pente ou de courbure (voir principe des **fonctions splines**). On peut aussi tout simplement construire des approximations globales ne passant pas exactement par tous les points d'appui (voir **méthode des moindres carrés**).

Ex : Les masses volumiques (ρ en kg/m^3) d'un échantillon de sodium sont mesurées pour trois températures (T en $^\circ\text{C}$) : $(T_0, \rho_0) = (94, 929)$, $(T_1, \rho_1) = (205, 902)$ et $(T_2, \rho_2) = (371, 860)$. Grâce à la formule d'interpolation de Lagrange, calculer la masse volumique du sodium pour $T = 251^\circ\text{C}$? $T = 305^\circ\text{C}$?

$$\rho(T) = \rho_0 \frac{(T-T_1)(T-T_2)}{(T_0-T_1)(T_0-T_2)} + \rho_1 \frac{(T-T_0)(T-T_2)}{(T_1-T_0)(T_1-T_2)} + \rho_2 \frac{(T-T_0)(T-T_1)}{(T_2-T_0)(T_2-T_1)}$$

$$= \frac{929}{30747}(T-205)(T-371) - \frac{902}{18426}(T-94)(T-371) + \frac{860}{45982}(T-94)(T-205)$$

On trouve $\rho = 890.56 \text{ kg/m}^3$ pour $T = 251^\circ\text{C}$ et $\rho = 876.93 \text{ kg/m}^3$ pour $T = 305^\circ\text{C}$.

Ex : Déterminer le polynôme d'interpolation de Lagrange satisfaisant au tableau de valeurs suivant.

x	0	2	3	5
y	-1	2	9	87

Comme il y a quatre points d'appui, le polynôme sera de degré 3.

$$A_0 = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} = \frac{(x-2)(x-3)(x-5)}{(-2)(-3)(-5)} = -\frac{1}{30}(x^3 - 10x^2 + 31x - 30)$$

$$A_1 = \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} = \frac{x(x-3)(x-5)}{(2)(-1)(-3)} = \frac{1}{6}(x^3 - 8x^2 + 15x)$$

$$A_2 = \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} = \frac{x(x-2)(x-5)}{(3)(1)(-2)} = -\frac{1}{6}(x^3 - 7x^2 + 10x)$$

$$A_3 = \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} = \frac{x(x-2)(x-3)}{(5)(3)(2)} = \frac{1}{30}(x^3 - 5x^2 + 6x)$$

$$\text{Donc } P(x) = y_0 A_0(x) + y_1 A_1(x) + y_2 A_2(x) + y_3 A_3(x) = \frac{53}{30}x^3 - 7x^2 + \frac{253}{30}x - 1$$

II.2. Application à l'intégration numérique

On vient de déterminer la fonction $f(x)$ passant exactement par tous les points expérimentaux $M_i(x_i, y_i)$, $i \in \{0, 1, \dots, n\}$. L'intégrale de cette fonction nous permet d'effectuer une **intégration numérique** de la courbe expérimentale. On supposera que les points sont équidistants, c'est-à-dire que $x_{i+1} - x_i = \Delta x$.

a) Formule des trapèzes

Considérons les deux premiers points $M_0(x_0, y_0)$ et $M_1(x_1, y_1)$. Nous avons vu que l'on pouvait joindre ces deux points par une droite d'équation : $P(x) = y_1 + (y_1 - y_0) \frac{x - x_0}{\Delta x}$. On obtient ainsi pour l'intégrale :

$$\int_{x_0}^{x_1} P(x) dx = \int_{x_0}^{x_1} \left(y_1 + (y_1 - y_0) \frac{x - x_0}{\Delta x} \right) dx = y_1 [x]_{x_0}^{x_1} + \frac{y_1 - y_0}{\Delta x} \left[\frac{x^2}{2} - x_0 x \right]_{x_0}^{x_1} = \frac{1}{2} \Delta x (y_0 + y_1)$$

Si l'on répète ce procédé pour tous les intervalles, on obtient la **formule des trapèzes** :

$$\int_{x_0}^{x_n} P(x) dx = \int_{x_0}^{x_1} P(x) dx + \int_{x_1}^{x_2} P(x) dx + \dots + \int_{x_{n-1}}^{x_n} P(x) dx = \frac{1}{2} \Delta x (y_0 + y_1) + \frac{1}{2} \Delta x (y_1 + y_2) + \dots + \frac{1}{2} \Delta x (y_{n-1} + y_n)$$

$$= \frac{1}{2} \Delta x (y_0 + 2y_1 + 2y_2 + \dots + 2y_{n-1} + y_n)$$

b) Formule de Simpson

Considérons maintenant les trois premiers points $M_0(x_0, y_0)$, $M_1(x_1, y_1)$ et $M_2(x_2, y_2)$. Nous avons vu que l'on pouvait rejoindre ces trois points par une parabole d'équation : $P(x) = y_1 + \frac{y_2 - y_0}{2} \frac{x - x_0}{\Delta x} + \frac{y_2 - 2y_1 + y_0}{2} \frac{(x - x_0)^2}{(\Delta x)^2}$. On obtient ainsi pour l'intégrale :

$$\begin{aligned}
\int_{x_0}^{x_2} P(x) dx &= \int_{x_0}^{x_2} \left(y_1 + \frac{y_2 - y_0}{2} \frac{x - x_1}{\Delta x} + \frac{y_2 - 2y_1 + y_0}{2} \frac{(x - x_1)^2}{(\Delta x)^2} \right) dx \\
&= y_1 [x]_{x_0}^{x_2} + \frac{y_2 - y_0}{2\Delta x} \left[\frac{x^2}{2} - x_1 x \right]_{x_0}^{x_2} + \frac{y_2 - 2y_1 + y_0}{6(\Delta x)^2} [(x - x_1)^3]_{x_0}^{x_2} \\
&= 2\Delta x y_1 + \frac{y_2 - 2y_1 + y_0}{3} \Delta x \\
&= \frac{1}{3} \Delta x (y_0 + 4y_1 + y_2)
\end{aligned}$$

Comme précédemment, en répétant ce procédé sur un nombre pair d'intervalles (c'est-à-dire un nombre impair de points), on obtient la **formule de Simpson** :

$$\begin{aligned}
\int_{x_0}^{x_n} P(x) dx &= \int_{x_0}^{x_2} P(x) dx + \int_{x_2}^{x_4} P(x) dx + \dots + \int_{x_{n-2}}^{x_n} P(x) dx \\
&= \frac{1}{3} \Delta x (y_0 + 4y_1 + y_2) + \frac{1}{3} \Delta x (y_2 + 4y_3 + y_4) + \dots + \frac{1}{3} \Delta x (y_{n-2} + 4y_{n-1} + y_n) \\
&= \frac{1}{3} \Delta x (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{n-2} + 4y_{n-1} + y_n)
\end{aligned}$$

Ex : Déterminer l'intégrale numérique de la fonction *sin* évaluée en cinq points par formule des trapèzes. Par la formule de Simpson.

x	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$
y	0	$\sqrt{2-\sqrt{2}}/2$	$\sqrt{2}/2$	$\sqrt{2+\sqrt{2}}/2$	1

car $\sin \frac{\alpha}{2} = \sqrt{\frac{1 - \cos \alpha}{2}}$ si $\alpha \in [0, \pi/2]$

Analytique : $I = \int_0^{\pi/2} \sin x dx = [-\cos x]_0^{\pi/2} = 1$

Trapèzes : $I = \frac{1}{2} \frac{\pi}{8} \left(0 + 2 \times \frac{\sqrt{2-\sqrt{2}}}{2} + 2 \times \frac{\sqrt{2}}{2} + 2 \times \frac{\sqrt{2+\sqrt{2}}}{2} + 1 \right) = 0.987116$

Simpson : $I = \frac{1}{3} \frac{\pi}{8} \left(0 + 4 \times \frac{\sqrt{2-\sqrt{2}}}{2} + 2 \times \frac{\sqrt{2}}{2} + 4 \times \frac{\sqrt{2+\sqrt{2}}}{2} + 1 \right) = 1.000135$

On constate que l'approximation de Simpson est nettement meilleure que celle fournie par la méthode des trapèzes.

Ex : Une voiture démarre et l'on mesure son accélération pendant les 4 premières secondes.

t	0	1	2	3	4
a	3.0	5.1	5.5	6.0	3.2

1) A l'aide de la formule de Simpson, déterminer la vitesse de la voiture à $t = 4$ s.

On a un nombre impair de points de mesures et un nombre pair d'intervalles. L'accélération et la vitesse sont liées par la relation :

$$a(t) = \frac{dv(t)}{dt} \Rightarrow v_4 = \int_0^4 a(t) dt = \frac{1}{3} \Delta t (a_0 + 4a_1 + 2a_2 + 4a_3 + a_4) = \frac{1}{3} \times 1 \times (3 + 4 \times 5.1 + 2 \times 5.5 + 4 \times 6 + 3.2)$$

Donc $v_4 = 20.53 \text{ m s}^{-1}$

2) Quelle est la distance parcourue par la voiture à $t = 4$ s ?

La vitesse et la distance sont liées par la relation :

$$v(t) = \frac{dx(t)}{dt}$$

Il faut donc calculer la vitesse à chaque temps. Cependant, on ne peut le faire avec la formule de Simpson que pour un nombre de points impair, c'est-à-dire pour $i = 4, 2$ et 0 . v_4 a été calculé à la question précédente.

$$v_2 = \int_0^2 a(t) dt = \frac{1}{3} \Delta t (a_0 + 4a_1 + a_2) = \frac{1}{3} \times 1 \times (3 + 4 \times 5.1 + 5.5) = 9.63 \text{ m s}^{-1}$$

$$v_0 = \int_0^0 a(t) dt = 0 \text{ m s}^{-1}$$

On obtient donc le tableau de valeurs suivant :

t	0	2	4
v	0.0	9.63	20.53

$$v(t) = \frac{dx(t)}{dt} \Rightarrow x_4 = \int_0^4 v(t) dt = \frac{1}{3} \Delta t (v_0 + 4v_1 + v_2) = \frac{1}{3} \times 2 \times (0 + 4 \times 9.63 + 20.53) = 39.37 \text{ m}$$

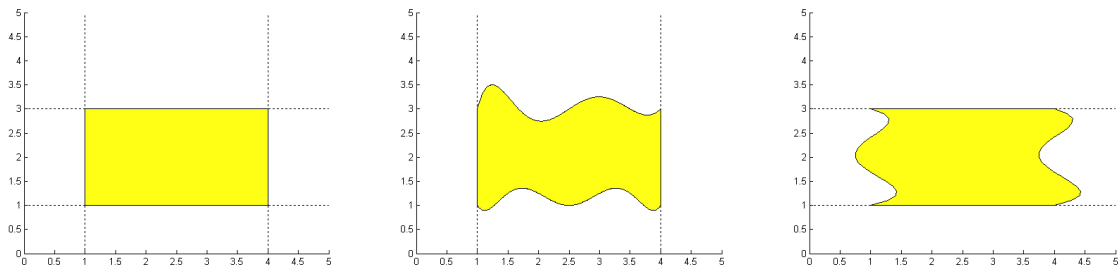
c) Rappels sur les intégrales doubles

Une intégrale double est une fonction du domaine de définition D , partie de \mathbb{R}^2 , vers \mathbb{R} : $\iint_D f(x, y) dx dy$.

Si le domaine est de la forme $D = [a, b] \times [c, d]$, alors $\iint_D f(x, y) dx dy = \int_a^b \left(\int_c^d f(y, y) dy \right) dx$

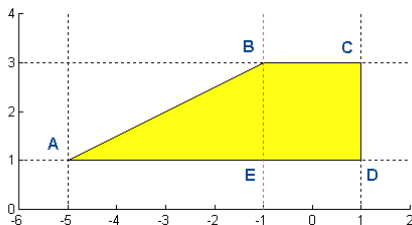
Si le domaine est de la forme $D = [a, b] \times [\varphi_1(x), \varphi_2(x)]$, alors $\iint_D f(x, y) dx dy = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(y, y) dy \right) dx$

Si le domaine est de la forme $D = [\psi_1(y), \psi_2(y)] \times [c, d]$, alors $\iint_D f(x, y) dx dy = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(y, y) dx \right) dy$



Si $D = D_1 \cup D_2$ et $D_1 \cap D_2 = \emptyset$ alors $\iint_D f(x, y) dx dy = \iint_{D_1} f(x, y) dx dy + \iint_{D_2} f(x, y) dx dy$

Ex : 1) Calculer l'intégrale $I = \iint_D (1+x)(1+y) dx dy$ analytiquement sur le domaine D .



On divise le domaine d'intégration D en deux sous domaines : D_1 constitué par le triangle (ABE) et D_2 constitué par le rectangle (BCDE). Considérons I_1 correspondant à D_1 . La droite (AB) a pour équation :

$$y = 1 \times \frac{x+1}{-4} + 3 \times \frac{x+5}{4} = \frac{1}{2}x + \frac{7}{2}$$

$$\text{Alors } I_1 = \int_{x=-5}^{x=-1} \int_{y=1}^{y=\frac{1}{2}x+\frac{7}{2}} (1+x)(1+y) dx dy$$

$$\begin{aligned} I_1 &= \int_{-5}^{-1} (1+x) \left(\int_1^{\frac{1}{2}x+\frac{7}{2}} (1+y) dy \right) dx = \int_{-5}^{-1} (1+x) \left[y + \frac{1}{2}y^2 \right]_1^{\frac{1}{2}x+\frac{7}{2}} dx = \frac{1}{8} \int_{-5}^{-1} (1+x)(x^2 + 18x + 65) dx \\ &= \frac{1}{8} \int_{-5}^{-1} (x^3 + 19x^2 + 83x + 65) dx = \frac{1}{8} \left[\frac{1}{4}x^4 + \frac{19}{3}x^3 + \frac{83}{2}x^2 + 65x \right]_{-5}^{-1} = -\frac{40}{3} \end{aligned}$$

Remarque : on aurait pu considérer que $y = \frac{1}{2}x + \frac{7}{2} \Leftrightarrow x = 2y - 7$ et calculer l'intégrale :

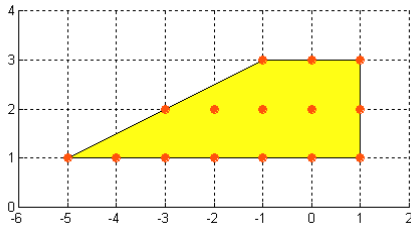
$$I_1 = \int_{y=1}^{y=3} \int_{x=2y-7}^{x=-1} (1+x)(1+y) dx dy = -\frac{40}{3}$$

De la même façon, on peut déterminer I_2 correspondant à D_2 : $I_2 = \int_{x=-1}^{x=1} \int_{y=1}^{y=3} (1+x)(1+y) dx dy$

$$I_2 = \int_{-1}^1 (1+x) \left(\int_1^3 (1+y) dy \right) dx = \int_{-1}^1 (1+x) \left[y + \frac{1}{2}y^2 \right]_1^3 dx = 6 \int_{-1}^1 (1+x) dx = 6 \left[x + \frac{1}{2}x^2 \right]_{-1}^1 = 12$$

$$\text{Donc } I = I_1 + I_2 = -\frac{40}{3} + 12 = -\frac{4}{3}$$

2) Refaire le calcul à l'aide de la formule de Simpson en posant $\Delta x = \Delta y = 1$.



On échantillonne le domaine D en un nombre impair de points en x et en y , et on applique la formule de Simpson dans les deux dimensions. Considérons les lignes de points correspondant à $y = 1$, $y = 2$, $y = 3$ et la fonction $f(x, y) = (1+x)(1+y)$

$$I_0 = \frac{1}{3} \Delta x (f(-5, 1) + 4f(-4, 1) + 2f(-3, 1) + 4f(-2, 1) + 2f(-1, 1) + 4f(0, 1) + f(1, 1)) = -12$$

$$I_1 = \frac{1}{3} \Delta x (f(-3, 2) + 4f(-2, 2) + 2f(-1, 2) + 4f(0, 2) + f(1, 2)) = 0$$

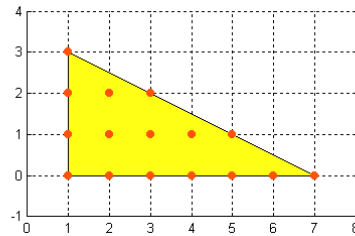
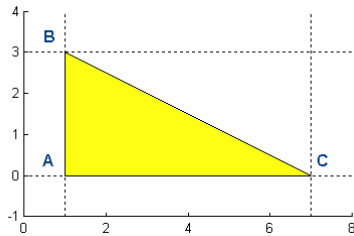
$$I_2 = \frac{1}{3} \Delta x (f(-1, 3) + 4f(0, 3) + f(1, 3)) = 8$$

On applique alors la formule de Simpson pour ces trois valeurs et on obtient l'intégrale :

$$I = \frac{1}{3} \Delta y (I_0 + 4I_1 + I_2) = \frac{1}{3} (-12 + 0 + 8) = -\frac{4}{3}$$

On retrouve bien le même résultat qu'avec la solution analytique.

Ex : 1) Calculer l'intégrale $I = \iint_D xy dx dy$ sur le domaine D par les méthodes analytique et numérique.



La droite (BC) a pour équation : $y = 3 \times \frac{x-7}{-6} + 0 \times \frac{x-1}{6} = -\frac{1}{2}x + \frac{7}{2}$. Alors :

$$I = \int_{x=1}^{x=7} \int_{y=0}^{y=-\frac{1}{2}x+\frac{7}{2}} xy dx dy I_1 = \int_1^7 x \left(\int_0^{-\frac{1}{2}x+\frac{7}{2}} y dy \right) dx = \int_1^7 x \left[\frac{1}{2} y^2 \right]_0^{-\frac{1}{2}x+\frac{7}{2}} dx = \frac{1}{8} \int_1^7 x (x^2 - 14x + 49) dx$$

$$= \frac{1}{8} \int_1^7 (x^3 - 14x^2 + 49x) dx = \frac{1}{8} \left[\frac{1}{4} x^4 - \frac{14}{3} x^3 + \frac{49}{2} x^2 \right]_1^7 = \frac{45}{2} = 22.5$$

Considérons les lignes de points correspondant à $y = 0$, $y = 1$ et $y = 2$ (la ligne $y = 3$ est provisoirement écartée car la méthode de Simpson requiert un nombre impair de points et donc un nombre pair d'intervalles) et la fonction $f(x, y) = xy$:

$$I_0 = \frac{1}{3} \Delta x (f(1, 0) + 4f(2, 0) + 2f(3, 0) + 4f(4, 0) + 2f(5, 0) + 4f(6, 0) + f(7, 0)) = 0$$

$$I_1 = \frac{1}{3} \Delta x (f(1, 1) + 4f(2, 1) + 2f(3, 1) + 4f(4, 1) + f(5, 1)) = 12$$

$$I_2 = \frac{1}{3} \Delta x (f(1, 2) + 4f(2, 2) + f(3, 2)) = 8$$

On applique alors la formule de Simpson pour ces trois valeurs et on obtient l'intégrale :

$$I_{012} = \frac{1}{3} \Delta y (I_0 + 4I_1 + I_2) = \frac{1}{3} (0 + 48 + 8) = \frac{56}{3}$$

L'intégrale de $f(x, y)$ sur le triangle délimité par les droites $y = 2$ et $y = 3$ est calculée par la méthode des trapèzes :

$$I_{23} = \frac{1}{2} \Delta y (I_2 + I_3) = \frac{1}{2} (8 + 3) = \frac{11}{2}$$

Donc au total $I = I_{012} + I_{23} = \frac{56}{3} + \frac{11}{2} = \frac{145}{6} = 24.17$. Dans cet exercice, la méthode numérique surestime la valeur de l'intégrale car l'approximation des trapèzes est moins précise que celle de Simpson.

III. Résolution d'équations par des méthodes itératives

En sciences pour l'ingénieur, il arrive très souvent que l'on veuille déterminer les racines d'une équation, c'est-à-dire la (ou les) valeur(s) de x telle(s) que $f(x) = 0$. Le calcul du point d'intersection entre deux courbes représentées par les fonctions $f(x)$ et $g(x)$ revient aussi à déterminer les **racines** de $h(x) = f(x) - g(x)$. Les techniques de résolution d'équations dépendent beaucoup des variations de la fonction et sont en général itératives.

III.1. Racines d'un polynôme

Soit $P_n(x)$ un polynôme de degré n : $P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n$. $P_n(x)$ admet exactement n racines réelle ou complexes, que l'on notera $r_1, r_2, \dots, r_{n-1}, r_n$. La recherche de ces racines se construit de la manière suivante : si l'on obtient une première racine, on peut écrire $P_n(x) = (x - r_1)P_{n-1}(x)$ où $P_{n-1}(x)$ est un polynôme de degré $n-1$. On recommence jusqu'à mettre le polynôme sous la forme $P_n(x) = a_n(x - r_1)(x - r_2)\dots(x - r_n) = a_n \prod_{i=1}^n (x - r_i)$. C'est la **méthode de réduction polynomiale**.

- **Théorème d'Abel** : il n'existe pas de formule explicite donnant les racines de $P_n(x)$ pour $n \geq 5$.
- **Règle des signes de Descartes** : elle permet de connaître le nombre de racines positives ou négatives. On compte v le nombre de changements de signe des coefficients du polynôme $a_0, a_1, a_2, \dots, a_n$. Le nombre de racines réelles positives est égal à v ou à un nombre pair inférieur à v ; le nombre de racines réelles négatives est déterminé en remplaçant x par $-x$ et en appliquant la même règle que pour les racines positives.

Ex : Déterminer le nombre de racines positives et négatives de la fonction $f(x) = x^4 - 5x^3 + 5x^2 + 5x - 6$.

$f(x)$ est un polynôme de degré 4 qui admet 4 racines réelles ou complexes. $nsc = 3 \rightarrow$ trois ou une racines réelles positives.

$f(-x) = x^4 + 5x^3 + 5x^2 - 5x - 6 \Rightarrow nsc = 1 \rightarrow$ une racine réelle négative.

Donc le polynôme admet 3 racines positives (1, 2, 3) et une négative (-1).

Il existe de nombreuses méthodes permettant de rechercher les zéros d'un polynôme. L'une d'entre elles consiste à construire la **matrice compagnon** du polynôme et à en calculer ses **valeurs propres** :

$$\text{Soit } P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n = 0 \Rightarrow x^n = -\frac{a_{n-1}}{a_n}x^{n-1} - \frac{a_{n-2}}{a_n}x^{n-2} - \dots - \frac{a_2}{a_n}x^2 - \frac{a_1}{a_n}x - \frac{a_0}{a_n}$$

La première ligne de la matrice compagnon, d'ordre n , reprend les coefficients du polynôme de degré $n-1$ du membre de droite, et s'écrit sous la forme :

$$A = \begin{pmatrix} -\frac{a_{n-1}}{a_n} & -\frac{a_{n-2}}{a_n} & \dots & -\frac{a_2}{a_n} & -\frac{a_1}{a_n} & -\frac{a_0}{a_n} \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

Les valeurs propres de cette matrice sont les racines du **polynôme caractéristique** de degré n défini par $\det(A - \lambda I) = 0$ où I désigne la matrice identité.

III.2. Racines d'une fonction quelconque

La plupart des fonctions ne s'expriment pas sous la forme de polynômes mais s'écrivent de façon plus ou moins complexe comme des combinaisons de fonctions de base : fonctions trigonométriques, fonctions exponentielle ou logarithme, etc. Le principe de recherche de racines d'équations est basé sur des **méthodes itératives** où, en partant d'une valeur initiale, on s'approche de plus en plus près de la solution exacte. Il est conseillé d'étudier les **variations de la fonction** et de la tracer sur un intervalle donné afin d'avoir une **idée préalable du voisinage de la racine**. La clé du succès repose sur l'encadrement préalable de la racine. La condition d'existence d'une solution (au moins) pour la fonction $f(x)$ peut s'énoncer de la façon suivante : si f est continue et s'il existe un intervalle $[a, b]$ tel que $f(a) \times f(b) \leq 0$ alors il existe au moins une racine réelle $r \in [a, b]$ telle que $f(r) = 0$.

a) Méthode de bisection ou de dichotomie

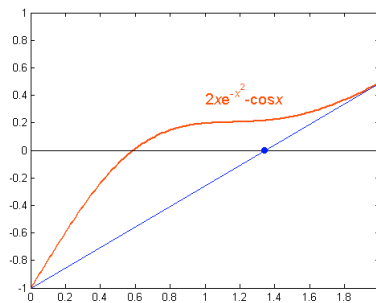
Supposons que l'on connaisse un intervalle $[x_0, x_1]$ contenant la racine r . Soient $y_0 = f(x_0)$ et $y_1 = f(x_1)$. On coupe l'intervalle en deux $x_m = \frac{1}{2}(x_0 + x_1)$ et on évalue $y_m = f(x_m)$. Soit $f(x_0) \times f(x_m) < 0$ et la racine se trouve sur l'intervalle $[x_0, x_m]$, soit $f(x_m) \times f(x_1) < 0$ et la racine se trouve sur l'intervalle $[x_m, x_1]$. On réitère cette opération jusqu'à ce qu'un **critère de convergence** soit atteint. Cette méthode est très populaire car la solution est garantie mais sa rapidité de convergence n'est pas bonne comparée à d'autres. Il est donc suggéré de ne l'utiliser que pour s'approcher de la racine.

b) Méthode de la fausse position

Il est possible d'améliorer la convergence de l'algorithme en interpolant la fonction par une droite sur l'intervalle $[x_0, x_1]$. Précédemment, nous avons établi l'équation de la droite reliant deux points en utilisant la méthode des polynômes de Lagrange :

$$P(x) = y_1 + (y_1 - y_0) \frac{x - x_1}{\Delta x} = y_1 + \frac{y_1 - y_0}{x_1 - x_0} (x - x_1)$$

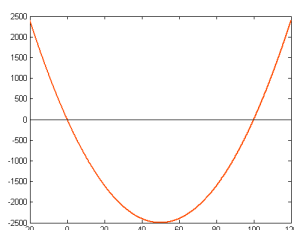
La nouvelle abscisse n'est plus le milieu de l'intervalle mais le point d'intersection entre la droite et l'axe des abscisses : $P(x_m) = 0 \Rightarrow y_1 + \frac{y_1 - y_0}{x_1 - x_0} (x_m - x_1) = 0 \Rightarrow x_m = x_1 - y_1 \frac{x_1 - x_0}{y_1 - y_0}$. On évalue $y_m = f(x_m)$ et on reprend à ce stade le principe de l'algorithme précédent. La relation de récurrence s'écrit : $x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$.



c) Méthode du point fixe

Elle est fondée sur le fait qu'il est toujours possible de transformer le problème $f(x) = 0$ en un problème équivalent $g(x) - x = 0$ où la **fonction auxiliaire** g a été choisie de manière à ce que $g(\alpha) = \alpha$ quand $f(\alpha) = 0$. Le choix de la fonction g n'est pas unique. La méthode revient à déterminer son intersection avec la première bissectrice des axes d'équation $y = x$. On part d'une abscisse initiale x_0 prise proche de la racine r cherchée puis on calcule $g(x_0) = x_1$; connaissant x_1 on calcule $g(x_1) = x_2$ et ainsi de suite jusqu'à obtenir une stabilisation du processus, c'est-à-dire $g(x_*) = x_*$ avec une précision donnée. La racine cherchée n'est alors autre que $r \approx x_*$. Cette méthode n'est pas toujours utilisable, elle doit satisfaire deux critères de convergence : si $x \in [a, b]$ alors $g(x) \in [a, b]$ et $-1 \leq g'(x) \leq 1$.

Ex : Déterminer les deux racines de la fonction $f(x) = x^2 - 100x + 1 = 0$ par la méthode du point fixe.



La méthode du point fixe nécessite d'écrire l'équation sous la forme $x = g(x)$

$$x^2 - 100x + 1 = 0 \Rightarrow x = \frac{1}{100}(x^2 + 1) \Rightarrow g(x) = \frac{1}{100}(x^2 + 1)$$

Calcul de la première racine sur l'intervalle $[-20, 20]$: $f(-20) = 2401$ et $f(20) = -1599$ donc $f(-20) \times f(20) < 0$ ce qui signifie qu'elle existe bien.

$$g(-20) = \frac{401}{100} = 4.01 \in [-20, 20] \text{ et } g(20) = \frac{401}{100} = 4.01 \in [-20, 20]$$

$g'(x) = \frac{x}{50} \Rightarrow -1 \leq g'(x) \leq 1$ sur l'intervalle $[-20, 20]$. Les deux critères de convergence étant vérifiés, la

méthode du point fixe peut être appliquée. On part de $x_0 = 1$ et on itère :

$$x_0 = 1 \Rightarrow x_1 = g(x_0) = 0.02, \quad x_1 = 0.02 \Rightarrow x_2 = g(x_1) = 0.010004, \quad x_2 = 0.010004 \Rightarrow x_3 = g(x_2) = 0.0100010008$$

$$x_3 = 0.010004 \Rightarrow x_4 = g(x_3) = 0.01000100020017$$

Calcul de la deuxième racine sur l'intervalle $[20, 120]$: $f(20) = -1599$ et $f(120) = 2401$ donc

$f(20) \times f(120) < 0$ ce qui signifie qu'elle existe bien.

$g(20) = \frac{401}{100} = 4.01 \notin [20, 120]$ donc le premier critère de convergence n'est pas satisfait... Il faut alors changer

d'écriture en posant $x^2 - 100x + 1 = 0 \Rightarrow x^2 = 100x - 1 \Rightarrow x = 100 - \frac{1}{x} \Rightarrow h(x) = 100 - \frac{1}{x}$

$$h(20) = \frac{1999}{20} = 99.95 \in [20, 120] \quad \text{et} \quad h(120) = \frac{11999}{120} = 99.99 \in [20, 120]$$

$h'(x) = \frac{1}{x^2} \Rightarrow -1 \leq h'(x) \leq 1$ sur l'intervalle $[20, 120]$. Les deux critères de convergence étant vérifiés, la

méthode du point fixe peut être appliquée. On part de $x_0 = 60$ et on itère :

$$x_0 = 60 \Rightarrow x_1 = h(x_0) = 99.98333333333333, \quad x_1 = 99.98333333333333 \Rightarrow x_2 = g(x_1) = 99.98999833305551,$$

$$x_2 = 99.98999833305551 \Rightarrow x_3 = g(x_2) = 99.98999899973326$$

d) Méthode des gradients ou de Newton-Raphson

Les méthodes précédentes ne nécessitent que la connaissance de la fonction en différents points de l'intervalle encadrant la racine. La méthode de Newton-Raphson est plus générale puisque l'on ne part que d'un point initial x_0 , proche de la racine, et non d'un intervalle $[x_0, x_1]$ qui l'encadre comme dans les deux premières méthodes. On écrit le développement de Taylor de f à l'ordre n en x_0 :

$$f(x) = f(x_0) + \frac{\partial f}{\partial x}(x_0) \frac{(x-x_0)}{1!} + \frac{\partial^2 f}{\partial x^2}(x_0) \frac{(x-x_0)^2}{2!} + \dots + \frac{\partial^n f}{\partial x^n}(x_0) \frac{(x-x_0)^n}{n!}$$

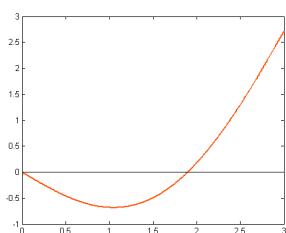
On cherche la valeur de x telle que $f(x) = 0 \Rightarrow f(x_0) + \frac{\partial f}{\partial x}(x_0)(x-x_0) \approx 0 \Rightarrow x \approx x_0 - \frac{f(x_0)}{f'(x_0)}$

Si x est proche de x_0 , c'est-à-dire si $|x-x_0|$ est faible, cette relation est vraie et on a une bonne approximation de la racine de la fonction. Sinon, le remplacement de la fonction $f(x)$ par les deux premiers termes de son développement de Taylor équivaut à approximer la fonction par sa tangente au point (x_0, y_0) . En écrivant que $f(x) = 0$, on calcule donc le point d'intersection entre la tangente et l'axe des abscisses. En conséquence, la valeur de l'abscisse ainsi générée n'est pas la solution mais une valeur plus proche que celle de l'abscisse du point initial. Il suffit donc de répéter

l'opération par le processus itératif $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$ jusqu'à convergence. Comme pour la méthode du point fixe, un

des critères de convergence est : $-1 < \left| \frac{f(x) f''(x)}{f'(x)^2} \right| < 1$.

Ex : Déterminer la racine de la fonction $f(x) = x - 2 \sin x = 0$ par la méthode de Newton-Raphson.



0 est une racine évidente de cette fonction.

$$f(x) = x - 2 \sin x \Rightarrow f'(x) = 1 - 2 \cos x$$

La relation de récurrence s'écrit : $x_{k+1} = x_k - \frac{x_k - 2 \sin x_k}{1 - 2 \cos x_k}$

On part de la valeur initiale $x_0 = 2$ et on itère :

$$x_1 = x_0 - \frac{x_0 - 2 \sin x_0}{1 - 2 \cos x_0} = 1.9010, \quad x_2 = x_1 - \frac{x_1 - 2 \sin x_1}{1 - 2 \cos x_1} = 1.8955 \quad \text{et les itérations}$$

suivantes donnent la même valeur. Si l'on part d'une valeur initiale plus éloignée de la solution, $x_0 = 1$ par exemple, on aboutit à la même solution mais avec un nombre plus grand d'itérations et des valeurs qui peuvent osciller : $x_1 = -7.4727$, $x_2 = 14.4785$, $x_3 = 6.9351$, $x_4 = 16.6357$, $x_5 = 8.3439$, $x_6 = 4.9546$, $x_7 = -8.3013$, $x_8 = -4.8173$, $x_9 = 3.7926$, $x_{10} = 1.8611$, $x_{11} = 1.8962$ et $x_{12} = 1.8955$

La méthode de Newton-Raphson permet généralement de trouver rapidement et avec une bonne précision les racines d'une fonction. Il existe cependant des situations où elle échoue, par exemple lorsque $f'(x_0)$ est faible, c'est-à-dire que la tangente est presque horizontale, à proximité d'un extrêum. Enfin son principal inconvénient est qu'elle nécessite l'évaluation de la fonction $f(x)$ et de sa dérivée $f'(x)$ qui n'est pas toujours facile à écrire analytiquement.

III.3. Systèmes non-linéaires

Nous nous sommes intéressés jusqu'à présent à des méthodes permettant de trouver le zéro d'une fonction d'une variable. On peut en principe les utiliser pour N fonctions des variables x_1, x_2, \dots, x_N . On se limitera ici à un système de deux équations non-linéaires à deux inconnues x et y :

$$\begin{cases} F(x, y) = 0 \\ G(x, y) = 0 \end{cases}$$

La méthode de Newton-Raphson est analogue à celle utilisée précédemment. On part d'un point initial (x_0, y_0) et on développe les deux fonctions F et G en série de Taylor à l'ordre 1 en x_0 et y_0 :

$$\begin{cases} F(x, y) = F(x_0, y_0) + (x - x_0) \frac{\partial F}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial F}{\partial y}(x_0, y_0) = 0 \\ G(x, y) = G(x_0, y_0) + (x - x_0) \frac{\partial G}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial G}{\partial y}(x_0, y_0) = 0 \end{cases}$$

qui se met facilement sous forme matricielle :

$$\begin{pmatrix} \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \\ \frac{\partial G}{\partial x}(x_0, y_0) & \frac{\partial G}{\partial y}(x_0, y_0) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = - \begin{pmatrix} F(x_0, y_0) \\ G(x_0, y_0) \end{pmatrix} \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \\ \frac{\partial G}{\partial x}(x_0, y_0) & \frac{\partial G}{\partial y}(x_0, y_0) \end{pmatrix}^{-1} \times \begin{pmatrix} F(x_0, y_0) \\ G(x_0, y_0) \end{pmatrix}$$

On peut écrire la formule de récurrence de Newton-Raphson pour deux inconnues :

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \begin{pmatrix} \frac{\partial F}{\partial x}(x_k, y_k) & \frac{\partial F}{\partial y}(x_k, y_k) \\ \frac{\partial G}{\partial x}(x_k, y_k) & \frac{\partial G}{\partial y}(x_k, y_k) \end{pmatrix}^{-1} \times \begin{pmatrix} F(x_k, y_k) \\ G(x_k, y_k) \end{pmatrix}$$

Ex : Trouver l'extremum de la fonction $f(x, y) = x^2 + 2xy + 2y^2 - 3x - 4y$ par la méthode de Newton-Raphson.

S'agit-il d'un maximum ou d'un minimum ?

$$\begin{cases} F(x, y) = \frac{\partial f}{\partial x}(x, y) = 2x + 2y - 3 = 0 \Rightarrow \frac{\partial F}{\partial x}(x, y) = \frac{\partial^2 f}{\partial x^2}(x, y) = 2 \text{ et } \frac{\partial F}{\partial y}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y) = 2 \\ G(x, y) = \frac{\partial f}{\partial y}(x, y) = 2x + 4y - 4 = 0 \Rightarrow \frac{\partial G}{\partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y) = 2 \text{ et } \frac{\partial G}{\partial y}(x, y) = \frac{\partial^2 f}{\partial y^2}(x, y) = 4 \end{cases}$$

On part du point initial $(x_0, y_0) = (0.5, 1)$, par exemple, et on calcule (x_1, y_1) par la formule ci-dessous :

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \\ \frac{\partial G}{\partial x}(x_0, y_0) & \frac{\partial G}{\partial y}(x_0, y_0) \end{pmatrix}^{-1} \times \begin{pmatrix} F(x_0, y_0) \\ G(x_0, y_0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}^{-1} \times \begin{pmatrix} F(x_0, y_0) \\ G(x_0, y_0) \end{pmatrix}$$

$$\text{Première itération : } \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}^{-1} \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

$$\text{Deuxième itération : } \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix} - \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}^{-1} \times \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

On constate qu'il suffit d'une seule itération pour trouver la solution exacte. Cela provient du fait que les dérivées partielles de $f(x, y)$ sont des fonctions linéaires. La formule de récurrence fournit donc la solution exacte après une seule itération. Pour savoir de quel type d'extremum il s'agit, on calcule les valeurs propres de la matrice des dérivées secondes appelée **matrice Hessienne** en $(1, 0.5)$:

$$\det(H - \lambda I) = \begin{vmatrix} 2 - \lambda & 2 \\ 2 & 4 - \lambda \end{vmatrix} = (2 - \lambda)(4 - \lambda) - 4 = \lambda^2 - 6\lambda + 4 = (\lambda - 3 - \sqrt{5})(\lambda - 3 + \sqrt{5})$$

$\lambda_1 = 3 + \sqrt{5} > 0$ et $\lambda_2 = 3 - \sqrt{5} > 0$ donc on se situe sur un minimum (si les deux valeurs sont négatives, on se situe sur un maximum).

Ex : Trouver l'extremum de la fonction de Rosenbrock $f(x, y) = 10(y - x^2)^2 + (1 - x)^2$ par la méthode de Newton-Raphson. S'agit-il d'un maximum ou d'un minimum ?

$$\begin{cases} F(x, y) = \frac{\partial f(x, y)}{\partial x} = -40x(y - x^2) - 2(1 - x) = 0 \\ G(x, y) = \frac{\partial f(x, y)}{\partial y} = 20(y - x^2) = 0 \end{cases}$$

$$\Rightarrow \frac{\partial F(x, y)}{\partial x} = \frac{\partial^2 f(x, y)}{\partial x^2} = -40(y - 3x^2) + 2 \text{ et } \frac{\partial F(x, y)}{\partial y} = \frac{\partial^2 f(x, y)}{\partial x \partial y} = -40x$$

$$\Rightarrow \frac{\partial G(x, y)}{\partial x} = \frac{\partial^2 f(x, y)}{\partial x \partial y} = -40x \text{ et } \frac{\partial G(x, y)}{\partial y} = \frac{\partial^2 f(x, y)}{\partial y^2} = 20$$

On part du point initial $(x_0, y_0) = (-1, 2)$, par exemple, et on itère le système ci-dessous :

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \\ \frac{\partial G}{\partial x}(x_0, y_0) & \frac{\partial G}{\partial y}(x_0, y_0) \end{pmatrix}^{-1} \times \begin{pmatrix} F(x_0, y_0) \\ G(x_0, y_0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} -40(y_0 - 3x_0^2) + 2 & -40x_0 \\ -40x_0 & 20 \end{pmatrix}^{-1} \times \begin{pmatrix} F(x_0, y_0) \\ G(x_0, y_0) \end{pmatrix}$$

$$\text{Première itération : } \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} - \begin{pmatrix} 42 & 40 \\ 40 & 20 \end{pmatrix}^{-1} \times \begin{pmatrix} 36 \\ 20 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} - \frac{1}{760} \begin{pmatrix} 20 & -40 \\ -40 & 42 \end{pmatrix} \times \begin{pmatrix} 36 \\ 20 \end{pmatrix} = \begin{pmatrix} -21 \\ 19 \\ 23 \\ 19 \end{pmatrix} = \begin{pmatrix} -1.11 \\ 1.21 \end{pmatrix}$$

$$\text{Deuxième itération : } \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} -1.11 \\ 1.21 \end{pmatrix} - \begin{pmatrix} 100.17 & 44.21 \\ 44.21 & 20.00 \end{pmatrix}^{-1} \times \begin{pmatrix} -4.70 \\ -0.22 \end{pmatrix} = \begin{pmatrix} 0.62 \\ -2.59 \end{pmatrix}$$

$$\text{Troisième itération : } \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0.62 \\ -2.59 \end{pmatrix} - \begin{pmatrix} 151.36 & -24.72 \\ -24.72 & 20.00 \end{pmatrix}^{-1} \times \begin{pmatrix} 72.66 \\ -59.40 \end{pmatrix} = \begin{pmatrix} 0.62 \\ 0.39 \end{pmatrix}$$

$$\text{Quatrième itération : } \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0.62 \\ 0.39 \end{pmatrix} - \begin{pmatrix} 33.19 & -24.98 \\ -24.98 & 20.00 \end{pmatrix}^{-1} \times \begin{pmatrix} -0.750 \\ -0.008 \end{pmatrix} = \begin{pmatrix} 0.9997 \\ 0.8686 \end{pmatrix}$$

$$\text{Cinquième itération : } \begin{pmatrix} x_5 \\ y_5 \end{pmatrix} = \begin{pmatrix} 0.9997 \\ 0.8686 \end{pmatrix} - \begin{pmatrix} 87.59 & -39.99 \\ -39.99 & 20.00 \end{pmatrix}^{-1} \times \begin{pmatrix} 5.63 \\ -2.82 \end{pmatrix} = \begin{pmatrix} 0.9998 \\ 0.9996 \end{pmatrix}$$

$$\text{Sixième itération : } \begin{pmatrix} x_6 \\ y_6 \end{pmatrix} = \begin{pmatrix} 0.9998 \\ 0.9996 \end{pmatrix} - \begin{pmatrix} 81.96 & -39.99 \\ -39.99 & 20.00 \end{pmatrix}^{-1} \times \begin{pmatrix} -0.0004426 \\ -0.0000001 \end{pmatrix} = \begin{pmatrix} 1.0000 \\ 1.0000 \end{pmatrix}$$

La solution est donc $(x, y) = (1, 1)$. Les valeurs propres de la matrice Hessienne en $(1, 1)$ sont :

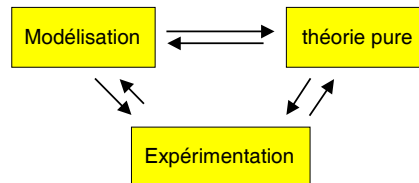
$$\det(H - \lambda I) = \begin{vmatrix} 82 - \lambda & -40 \\ -40 & 20 - \lambda \end{vmatrix} = (82 - \lambda)(20 - \lambda) - 1600 = \lambda^2 - 102\lambda + 40 = (\lambda - 51 - \sqrt{2561})(\lambda - 51 + \sqrt{2561})$$

$\lambda_1 = 51 + \sqrt{2561} > 0$ et $\lambda_2 = 51 - \sqrt{2561} > 0$ donc on se situe sur un minimum.

IV.1 La modélisation

a) Introduction

On appelle **modélisation** la **représentation mathématique simplifiée d'un phénomène permettant d'en simuler son fonctionnement**. Les sciences ont toujours utilisé, explicitement ou implicitement, la modélisation dans une démarche l'associant avec la **théorie pure** et l'**expérimentation**. Ces trois voies de recherche sont indissociables.



La **simulation informatique** d'un modèle a pour objet d'analyser les conséquences d'une théorie, de découvrir des effets nouveaux ou de prédire des comportements qualitatifs mais aussi de donner des réponses quantitatives :

- dans les **sciences exactes et technologiques**, elle dépend directement des moyens de calcul : l'augmentation continue de la puissance des ordinateurs permet d'imaginer des modèles de plus en plus riches et donc de s'attaquer à des phénomènes de plus en plus complexes. La science fondamentale en tire avantage par une meilleure compréhension des mécanismes en jeu, ouvrant de nouveaux champs d'étude comme celui des phénomènes non linéaires, du chaos dans les systèmes atomiques, de la réactivité en phase gazeuse, de la dynamique moléculaire, etc. L'industrie y gagne en temps et en coût, remplaçant l'étude expérimentale par la simulation dans des sujets comme l'aérodynamique ou la combustion dont l'importance est vitale pour l'aéronautique ou l'industrie automobile, aussi bien que pour prévoir l'effet thérapeutique d'une molécule.
- en **sciences humaines et sociales**, elle est d'avantage un guide de la réflexion, obligeant le chercheur à préciser sa problématique. Par exemple, l'instauration de systèmes d'information géographique (SIG) en vue de l'aménagement du territoire a nécessité une modélisation des objets graphiques et de leurs attributs. La tâche n'est pas toujours facile comme en témoigne la relative faiblesse de la modélisation en linguistique dont les querelles d'école et l'insuffisance de formation scientifique freinent le développement.

Le plus souvent le modèle emploie un **formalisme mathématique** et donc requiert des moyens de calcul importants. Aujourd'hui la puissance d'un **supercalculateur vectoriel** approche le **téraflop** (mille milliards d'opérations par seconde). En même temps que progressent les supercalculateurs, des machines plus petites mais atteignant déjà des performances remarquables, en général des **stations de travail**, s'introduisent dans les laboratoires. Elles permettent de tester des calculs à échelle réduite avant leur lancement sur une machine plus puissante et peuvent se connecter entre elles pour mettre leur puissance en commun. Enfin en bas de la chaîne les **micro-ordinateurs** commencent eux aussi à disposer d'une puissance de calcul intéressante.

b) Validation d'un modèle

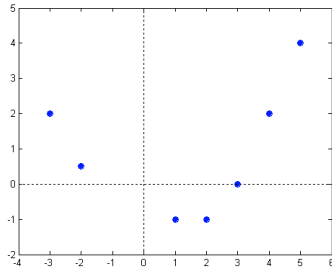
On attend d'une théorie qu'elle explique des résultats expérimentaux et qu'elle prédise de nouveaux résultats, alors qu'on attend généralement d'une expérience qu'elle vérifie la validité de théories existantes et qu'elle collecte des données pour les modifier. A l'exception de quelques cas où l'expérimentation est inaccessible, par exemple la création et la destruction des molécules interstellaires ou bien la rentrée de la navette spatiale dans l'atmosphère, la confrontation du modèle avec les **résultats expérimentaux** est nécessaire pour le **valider**. Il arrive cependant que cette confrontation soit trompeuse : ainsi, par pure coïncidence, une théorie fautive peut être en accord avec une expérience. Inversement une théorie peut être correcte mais en désaccord avec une expérience... De même il existe des "bonnes" et des "mauvaises" expériences. On supposera par la suite que les données expérimentales ont été calibrées, vérifiées, et débarrassées de leurs erreurs systématiques. Un phénomène physique peut être décrit par deux approches différentes :

1. **approche statistique** : on collecte des données expérimentales sur différents objets, et on essaie d'obtenir une corrélation entre les objets et les mesures. Cependant prédire n'est pas expliquer.
2. **approche analytique** : on essaie de comprendre les mécanismes du phénomène observé et on incorpore cette connaissance dans un modèle. C'est cette approche que nous développerons.

La complexité des modèles dépend de l'utilisation qu'on veut en faire. Dans la perspective d'une **inversion**, étape essentielle pour valider le modèle, le choix du modèle est gouverné par un certain nombre de règles. Il est clair qu'étant donné n'importe quel phénomène on peut toujours construire un modèle qui le décrit. En d'autres termes, si l'on utilise des paramètres en nombre suffisant, des fonctions d'un degré assez élevé, on peut toujours construire un modèle mathématique de n'importe quoi ! Mais le vrai problème n'est évidemment pas là : il consiste, en revanche, dans la

construction d'un modèle qui ne recoure pas à trop d'être mathématiques. Il y a donc un conflit entre l'adhérence rigoureuse à la donnée empirique, ce que les Anglo-Saxons appellent le *fit*, et la quantité de paramètres qui apparaissent dans le modèle : si on met beaucoup de paramètres, on a un bon *fit*, mais un modèle compliqué, si on en met peu, le modèle est simple, mais le *fit* est mauvais. Seront excellents les modèles qui concilient, moyennant quelques compromis, peu de paramètres et un bon *fit*.

c) Sens direct et sens inverse



Considérons un modèle représenté par une fonction mathématique f de p paramètres $\Theta = (\theta_1, \dots, \theta_p)$ indépendants. En **mode direct**, les paramètres sont fournis et l'on cherche à calculer un phénomène. Le **mode inverse** correspond à l'extraction des propriétés du milieu à partir des mesures physiques. Cette opération se fait par ajustement de la fonction mathématique f sur l'ensemble des points expérimentaux $M_i(x_i, y_i)$, $i \in \{1, \dots, n\}$. Soit e_i la distance entre les valeurs de y_i et $f(x_i)$: $e_i = y_i - f(x_i, (\theta_1, \dots, \theta_p))$. La **méthode des moindres carrés** consiste à chercher les valeurs des paramètres $\theta_1, \dots, \theta_p$ de f

en minimisant la **fonction de mérite** χ^2 définie comme la somme des carrés des distances e_i :

$$\chi^2(\theta_1, \dots, \theta_p) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n \left(y_i - f(x_i, (\theta_1, \dots, \theta_p)) \right)^2$$

En effet, certains écarts étant positifs et d'autres négatifs, on ne peut pas minimiser leur somme arithmétique. Le critère le plus couramment utilisé est donc la somme des carrés car ceux-ci sont toujours positifs et ne s'éliminent pas les uns les autres. Nous admettons ici une propriété des fonctions numériques à plusieurs variables : une condition nécessaire, mais pas suffisante, d'extremum en un point $\Theta = (\theta_1, \dots, \theta_p)$ est que toutes les **dérivées partielles** de χ^2 , à savoir

$\frac{\partial \chi^2}{\partial \theta_1}, \dots, \frac{\partial \chi^2}{\partial \theta_p}$, soient **nulles** en ce point.

IV.2. Ajustement à une fonction polynômiale

a) Polynôme de degré p

On pose $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_kx^k + \dots + a_px^p = \sum_{j=0}^p a_jx^j$ et on détermine les $p+1$ paramètres $a_0, a_1, a_2, \dots, a_p$

de cette fonction qui minimisent $\chi^2 = \sum_{i=1}^n \left(y_i - \sum_{j=0}^p a_jx_i^j \right)^2$. Pour tout k on écrit alors :

$$\frac{\partial \chi^2}{\partial a_k} = \frac{\partial}{\partial a_k} \left(\sum_{i=1}^n \left(y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_kx_i^k - \dots - a_px_i^p \right)^2 \right) = -2 \sum_{i=1}^n x_i^k \left(y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_kx_i^k - \dots - a_px_i^p \right)$$

En utilisant la propriété de dérivation suivante : $\frac{\partial}{\partial u} (\alpha - \beta u)^2 = -2\beta(\alpha - \beta u)$

$$\frac{\partial \chi^2}{\partial a_k} = 0 \Rightarrow a_0 \sum_{i=1}^n x_i^k + a_1 \sum_{i=1}^n x_i^{k+1} + a_2 \sum_{i=1}^n x_i^{k+2} + \dots + a_k \sum_{i=1}^n x_i^{2k} + \dots + a_p \sum_{i=1}^n x_i^{k+p} = \sum_{i=1}^n x_i^k y_i$$

En posant $S_k = \sum_{i=1}^n x_i^k$ et $W_k = \sum_{i=1}^n x_i^k y_i$ et en écrivant le système pour toutes les variables $a_0, a_1, a_2, \dots, a_p$ on obtient :

$$\begin{cases} na_0 + S_1 a_1 + S_2 a_2 + \dots + S_p a_p = W_0 \\ S_1 a_0 + S_2 a_1 + S_3 a_2 + \dots + S_{p+1} a_p = W_1 \\ S_2 a_0 + S_3 a_1 + S_4 a_2 + \dots + S_{p+2} a_p = W_2 \\ \dots \\ S_k a_0 + S_{k+1} a_1 + S_{k+2} a_2 + \dots + S_{p+k} a_p = W_k \\ \dots \\ S_p a_0 + S_{p+1} a_1 + S_{p+2} a_2 + \dots + S_{2p} a_p = W_p \end{cases} \Rightarrow \begin{pmatrix} n & S_1 & S_2 & \dots & S_p \\ S_1 & S_2 & S_3 & \dots & S_{p+1} \\ S_2 & S_3 & S_4 & \dots & S_{p+2} \\ \vdots & \vdots & \vdots & & \vdots \\ S_k & S_{k+1} & S_{k+2} & & S_{p+k} \\ \vdots & \vdots & \vdots & & \vdots \\ S_p & S_{p+1} & S_{p+2} & \dots & S_{2p} \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_k \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} W_0 \\ W_1 \\ W_2 \\ \vdots \\ W_k \\ \vdots \\ W_p \end{pmatrix} \Rightarrow SA = W$$

On constate que la matrice du système est une matrice carrée d'ordre $p+1$, symétrique, donc les $p+1$ paramètres sont obtenus en calculant l'inverse de S : $SA = W \Rightarrow S^{-1}(SA) = S^{-1}W \Rightarrow (S^{-1}S)A = S^{-1}W \Rightarrow IA = S^{-1}W \Rightarrow A = S^{-1}W$.

Ex : Ajustement par un polynôme du second degré (parabole) d'un nuage de 7 points : $M_1(-3,2)$,

$$M_2 = (-2, 0.5), M_3 = (1, -1), M_4 = (2, -1), M_5 = (3, 0), M_6 = (4, 2), M_7 = (5, 4)$$

Le polynôme s'écrit $f(x) = a_0 + a_1x + a_2x^2$ et ses coefficients sont calculés par $\begin{pmatrix} n & S_1 & S_2 \\ S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} W_0 \\ W_1 \\ W_2 \end{pmatrix}$

$$\text{avec } n = 7, S_1 = \sum_{i=1}^7 x_i = 10, S_2 = \sum_{i=1}^7 x_i^2 = 68, S_3 = \sum_{i=1}^7 x_i^3 = 190, S_4 = \sum_{i=1}^7 x_i^4 = 1076, W_0 = \sum_{i=1}^7 y_i = 6.5,$$

$$W_1 = \sum_{i=1}^7 x_i y_i = 18 \text{ et } W_2 = \sum_{i=1}^7 x_i^2 y_i = 147. \text{ Le système matriciel s'écrit donc :}$$

$$\begin{pmatrix} 7 & 10 & 68 \\ 10 & 68 & 190 \\ 68 & 190 & 1076 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 6.5 \\ 18 \\ 147 \end{pmatrix} \Rightarrow \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -1.2584 \\ -0.3043 \\ 0.2699 \end{pmatrix} \text{ et } f(x) = -1.2584 - 0.3043x + 0.2699x^2$$

Cas particulier de la droite de régression

La droite de régression s'écrit $f(x) = a_0 + a_1x$ avec a_0 l'ordonnée à l'origine et a_1 la pente de la droite. Ces deux coefficients qui sont les paramètres du modèle linéaire sont calculés par :

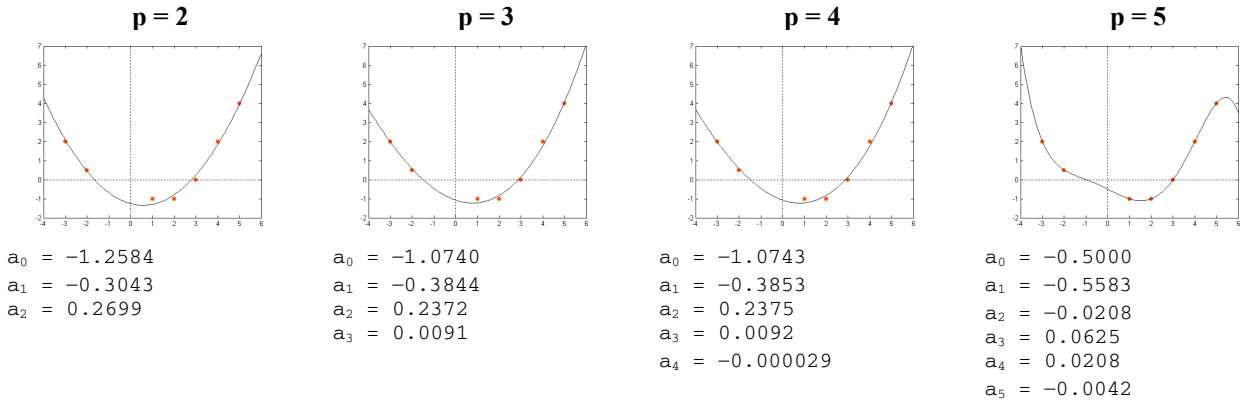
$$\begin{pmatrix} n & S_1 \\ S_1 & S_2 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} W_0 \\ W_1 \end{pmatrix} \Rightarrow \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} n & S_1 \\ S_1 & S_2 \end{pmatrix}^{-1} \times \begin{pmatrix} W_0 \\ W_1 \end{pmatrix} \Rightarrow \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{1}{nS_2 - (S_1)^2} \begin{pmatrix} S_2 & -S_1 \\ -S_1 & n \end{pmatrix}^{-1} \times \begin{pmatrix} W_0 \\ W_1 \end{pmatrix}$$

$$\text{donc } \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{1}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \begin{pmatrix} \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i \\ n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \end{pmatrix}$$

On retrouve les expressions bien connues pour a_0 et a_1 .

b) Généralisation

Les valeurs des coefficients $a_0, a_1, a_2, \dots, a_p$ du polynôme peuvent être très variables selon son degré, ce qui signifie qu'elles ne représentent pas les caractéristiques intrinsèques des données. Par exemple, la valeur a_0 est le point d'intersection sur l'axe des ordonnées, a_1 la pente en ce même point, et les autres coefficients les ordres supérieurs de courbure en ce même point.



Cette interdépendance serait évitée si l'on pouvait déterminer des coefficients représentant la valeur moyenne, la pente moyenne, la courbure moyenne, etc. des données. Ces derniers seraient alors indépendants du système de coordonnées. Une généralisation de l'approximation polynomiale consiste à déterminer le jeu de $p+1$ coefficients $a_0, a_1, a_2, \dots, a_p$ de la combinaison linéaire de fonctions $f_0(x), f_1(x), f_2(x), \dots, f_p(x)$:

$$f(x) = a_0 f_0(x) + a_1 f_1(x) + a_2 f_2(x) + \dots + a_k f_k(x) + \dots + a_p f_p(x)$$

On détermine les $p+1$ variables $a_0, a_1, a_2, \dots, a_p$ de cette fonction en minimisant $\chi^2 = \sum_{i=1}^n \left(y_i - \sum_{j=0}^p a_j f_j(x_i) \right)^2$. Pour tout k on écrit alors :

$$\begin{aligned} \frac{\partial \chi^2}{\partial a_k} &= \frac{\partial}{\partial a_k} \left(\sum_{i=1}^n (y_i - a_0 f_0(x_i) - a_1 f_1(x_i) - a_2 f_2(x_i) - \dots - a_k f_k(x_i) - \dots - a_p f_p(x_i))^2 \right) \\ &= -2 \sum_{i=1}^n f_k(x_i) (y_i - a_0 f_0(x_i) - a_1 f_1(x_i) - a_2 f_2(x_i) - \dots - a_k f_k(x_i) - \dots - a_p f_p(x_i)) \end{aligned}$$

Comme précédemment, on annule les dérivées partielles : $\frac{\partial \chi^2}{\partial a_k} = 0$

$$a_0 \sum_{i=1}^n f_0(x_i) f_k(x_i) + a_1 \sum_{i=1}^n f_1(x_i) f_k(x_i) + \dots + a_k \sum_{i=1}^n f_k(x_i)^2 + \dots + a_p \sum_{i=1}^n f_p(x_i) f_k(x_i) = \sum_{i=1}^n f_k(x_i) y_i$$

On obtient donc le système matriciel suivant que l'on sait résoudre :

$$\begin{pmatrix} \sum_{i=1}^n f_0(x_i)^2 & \sum_{i=1}^n f_0(x_i) f_1(x_i) & \sum_{i=1}^n f_0(x_i) f_2(x_i) & \dots & \sum_{i=1}^n f_0(x_i) f_p(x_i) \\ \sum_{i=1}^n f_1(x_i) f_0(x_i) & \sum_{i=1}^n f_1(x_i)^2 & \sum_{i=1}^n f_1(x_i) f_2(x_i) & \dots & \sum_{i=1}^n f_1(x_i) f_p(x_i) \\ \sum_{i=1}^n f_2(x_i) f_0(x_i) & \sum_{i=1}^n f_2(x_i) f_1(x_i) & \sum_{i=1}^n f_2(x_i)^2 & \dots & \sum_{i=1}^n f_2(x_i) f_p(x_i) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n f_p(x_i) f_0(x_i) & \sum_{i=1}^n f_p(x_i) f_1(x_i) & \sum_{i=1}^n f_p(x_i) f_2(x_i) & \dots & \sum_{i=1}^n f_p(x_i)^2 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n f_0(x_i) y_i \\ \sum_{i=1}^n f_1(x_i) y_i \\ \sum_{i=1}^n f_2(x_i) y_i \\ \vdots \\ \sum_{i=1}^n f_p(x_i) y_i \end{pmatrix}$$

Les fonctions f peuvent prendre des formes variées : si l'on parvenait à les choisir de façon à ce que les termes en dehors de la diagonale soient très faibles voire nuls, alors la matrice S serait diagonale et les coefficients $a_0, a_1, a_2, \dots, a_p$ indépendants du degré de la fonction. C'est le cas des **polynômes de Legendre**, des **polynômes de Tchebycheff**, etc. qui ont pour propriété d'être **orthogonaux**, c'est-à-dire que $\sum_{i=1}^n f_j(x_i) f_k(x_i) = 0$.

c) Ajustement à des fonctions orthogonales

- Polynômes de Legendre : ce sont des fonctions de x définies sur l'intervalle $[-1, 1]$ satisfaisant les propriétés suivantes :

- $P_p(x)$ est un polynôme en x de degré p qui contient seulement des puissances impaires de x si p est impair et seulement des puissances paires de x si p est pair.
- $P_p(1) = 1, \forall p \in \mathbb{N}$.
- $P_p(x)$ et $P_q(x)$ sont orthogonaux sur l'intervalle $[-1, 1]$

$$\int_{-1}^1 P_p(x) P_q(x) dx = \begin{cases} \frac{2}{2p+1} & \text{pour } p = q \\ 0 & \text{pour } p \neq q \end{cases}$$

Il existe plusieurs relations permettant de les calculer de proche en proche :

Relation de récurrence : $P_{p+1}(x) = \frac{2p+1}{p+1} x P_p(x) - \frac{p}{p+1} P_{p-1}(x)$

Formule de Rodrigues : $P_p(x) = \frac{1}{2^p p!} \frac{d^p}{dx^p} [(x^2 - 1)^p]$

Ex : Déterminer $P_0(x)$, $P_1(x)$, et $P_2(x)$

$P_0(x)$ est de la forme $P_0(x) = a_0$

$$\int_{-1}^1 P_0(x)^2 dx = 2 \Rightarrow a_0 = 1 \text{ donc } P_0(x) = 1$$

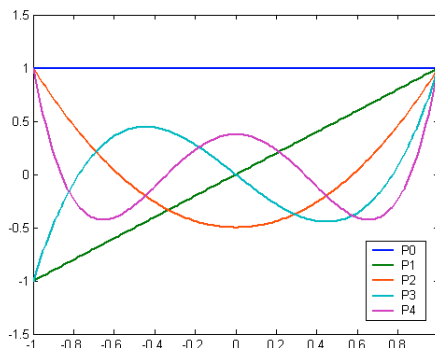
$P_1(x)$ est de la forme $P_1(x) = a_0 + a_1 x$

$$\left. \begin{aligned} \int_{-1}^1 P_0(x) P_1(x) dx = 0 &\Rightarrow a_0 = 0 \\ P_1(1) = 1 &\Rightarrow a_1 = 1 \end{aligned} \right\} \text{ donc } P_1(x) = x$$

$P_2(x)$ est de la forme $P_2(x) = a_0 + a_1 x + a_2 x^2$

$$\text{donc } P_2(x) = \frac{3}{2} x P_1(x) - \frac{1}{2} P_0(x) = \frac{1}{2} (3x^2 - 1)$$

On pourrait montrer que $P_3(x) = \frac{1}{2} (5x^3 - 3x)$, $P_4(x) = \frac{1}{8} (35x^4 - 30x^2 + 3)$, etc.



Ex : Ajustement par des polynômes de Legendre de degré 2 d'un nuage de 7 points : $(x_1, y_1) = (-3, 2)$,

$(x_2, y_2) = (-2, 0.5)$, $(x_3, y_3) = (1, -1)$, $(x_4, y_4) = (2, -1)$, $(x_5, y_5) = (3, 0)$, $(x_6, y_6) = (4, 2)$, $(x_7, y_7) = (5, 4)$

On fait d'abord un changement de variable afin que les abscisses soient comprises entre -1 et 1 . Les valeurs extrêmes de x étant $\min(x) = -3$ et $\max(x) = 5$, la nouvelle variable u est calculée par la formule :

$$u = \frac{1}{\max(x) - \min(x)} (2x - \max(x) - \min(x)) = \frac{1}{4} (x - 1)$$

$$u = \left[-1, -\frac{3}{4}, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1 \right] \text{ et le polynôme s'écrit } f(x) = a_0 P_0(u) + a_1 P_1(u) + a_2 P_2(u) \text{ avec } P_0(u) = 1,$$

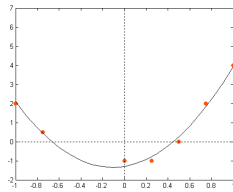
$P_1(u) = u$, $P_2(u) = \frac{1}{2} (3u^2 - 1)$. Ses coefficients sont calculés par :

$$\begin{pmatrix} \sum P_0(u_i)^2 & \sum P_0(u_i) P_1(u_i) & \sum P_0(u_i) P_2(u_i) \\ \sum P_1(u_i) P_0(u_i) & \sum P_1(u_i)^2 & \sum P_1(u_i) P_2(u_i) \\ \sum P_2(u_i) P_0(u_i) & \sum P_2(u_i) P_1(u_i) & \sum P_2(u_i)^2 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum P_0(u_i) y_i \\ \sum P_1(u_i) y_i \\ \sum P_2(u_i) y_i \end{pmatrix}$$

Le système matriciel s'écrit donc :

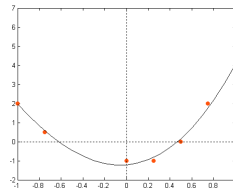
$$\begin{pmatrix} 7 & 3 & 53 \\ 3 & 55 & -21 \\ 53 & 21 & 2731 \\ 32 & 128 & 1024 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 13 \\ 23 \\ 497 \\ 64 \end{pmatrix} \Rightarrow \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 129 & -52 & -58 \\ 743 & -1205 & -525 \\ 52 & 85 & 44 \\ -1205 & 281 & 969 \\ 58 & 44 & 649 \\ -525 & 969 & 1454 \end{pmatrix} \times \begin{pmatrix} 13 \\ 23 \\ 497 \\ 64 \end{pmatrix} = \begin{pmatrix} 17 \\ 116 \\ 275 \\ 783 \\ 292 \\ 272 \end{pmatrix} = \begin{pmatrix} 0.1466 \\ 0.9418 \\ 2.8787 \end{pmatrix}$$

p = 2



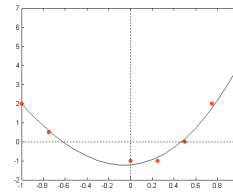
$$\begin{aligned} a_0 &= 0.1466 \\ a_1 &= 0.9418 \\ a_2 &= 2.8787 \end{aligned}$$

p = 3



$$\begin{aligned} a_0 &= 0.1989 \\ a_1 &= 0.8195 \\ a_2 &= 2.8219 \\ a_3 &= 0.2335 \end{aligned}$$

p = 4



$$\begin{aligned} a_0 &= 0.1991 \\ a_1 &= 0.8194 \\ a_2 &= 2.8228 \\ a_3 &= 0.2336 \\ a_4 &= -0.0017 \end{aligned}$$

Si l'on calcule $f(x)$ on retrouve un résultat obtenu dans un exercice précédent.

$$\begin{aligned} f(x) &= \frac{17}{116} \times P_0\left(\frac{1}{4}(x-1)\right) + \frac{275}{292} \times P_1\left(\frac{1}{4}(x-1)\right) + \frac{783}{272} \times P_2\left(\frac{1}{4}(x-1)\right) \\ &= \frac{17}{116} + \frac{275}{292} \times \frac{1}{4}(x-1) + \frac{783}{272} \times \frac{1}{2} \times \left(\frac{3}{16}(x-1)^2 - 1\right) \\ &= -\frac{23186927}{18426368} - \frac{96677}{317696}x + \frac{2349}{8704}x^2 \\ &= -1.2584 - 0.3043x + 0.2699x^2 \end{aligned}$$

- **Polynômes de Tchebycheff** : ce sont des fonctions de x définies sur l'intervalle $[-1,1]$ satisfaisant les propriétés suivantes :

1. $T_p(x)$ est un polynôme en x de degré p qui contient seulement des puissances impaires de x si p est impair et seulement des puissances paires de x si p est pair.
2. $T_p(1) = 1, \forall p \in \mathbb{N}$.
3. $T_p(x)$ et $T_q(x)$ sont orthogonaux sur l'intervalle $[-1,1]$, c'est-à-dire que :

$$\int_{-1}^1 \frac{T_p(x)T_q(x)}{\sqrt{1-x^2}} dx = \begin{cases} \pi & \text{pour } p = q = 0 \\ \frac{\pi}{2} & \text{pour } p = q \neq 0 \\ 0 & \text{pour } p \neq q \end{cases} \quad \text{ou} \quad \sum_{i=1}^n T_p(x_i)T_q(x_i) = \begin{cases} m & \text{pour } p = q = 0 \\ \frac{m}{2} & \text{pour } p = q \neq 0 \\ 0 & \text{pour } p \neq q \end{cases}$$

Les polynômes de Tchebycheff peuvent être calculés par la relation de récurrence $T_{p+1}(x) = 2xT_p(x) - T_{p-1}(x)$. Les premiers sont $T_0(x) = 1, T_1(x) = x, T_2(x) = 2x^2 - 1, T_3(x) = 4x^3 - 3x, T_4(x) = 8x^4 - 8x^2 + 1$, etc. La méthode est exactement la même que pour les polynômes de Legendre.

- **Séries de Fourier** : les fonctions sinus et cosinus sont aussi des fonctions orthogonales qui sont souvent utilisées pour modéliser des **séries temporelles** (*time series analysis*), par exemple des cycles horaires, journaliers, annuels de phénomènes naturels (température, éclairement solaire, flux de chaleur dans le sol, etc.) La fonction d'ajustement s'écrit $f(x) = \sum_{i=1}^{\infty} a_i \cos\left(\frac{2\pi x}{c_i}\right) + b_i \sin\left(\frac{2\pi x}{c_i}\right)$ où les termes a_i et b_i sont les amplitudes et les termes c_i sont les périodes des données.

d) **Analyse de surfaces polynomiales théoriques**

La régression polynomiale forme aussi la base de la méthode la plus simple d'analyse spatiale : l'analyse des surfaces théoriques (*trend surface analysis*). Cette technique consiste à modéliser la distribution spatiale d'une variable dépendante z à l'aide d'un polynôme de coordonnées x et y des observations. L'idée est d'estimer la valeur d'une variable sur la base de sa localisation. La technique est la même que celle de la régression polynomiale :

A l'ordre 1 : $f(x, y) = a_0 + a_1x + a_2y$

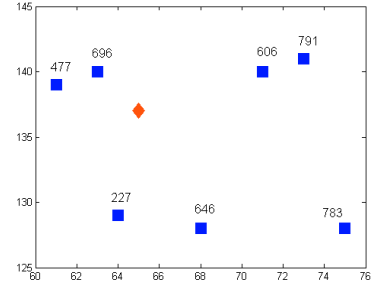
A l'ordre 2 : $f(x, y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2$

A l'ordre 3 : $f(x, y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + a_6y^3 + a_7x^2y + a_8xy^2 + a_9y^3$

On peut construire des polynômes d'ordre supérieur mais le nombre de paramètres devient rapidement très élevé, ce qui interdit l'usage de telles équations sur de petits échantillons. L'exemple suivant qui est une **interpolation spatiale** servira à illustrer notre propos.

Ex : La concentration en un polluant (z exprimée en ppm) a été mesurée en sept points de mesures localisés par leurs coordonnées spatiales x et y . Le but est d'estimer la concentration au point M_0 situé à la position $(65,137)$ par différentes méthodes.

	1	2	3	4	5	6	7
x_i	61	63	64	68	71	73	75
y_i	139	140	129	128	140	141	128
z_i	477	696	227	646	606	791	783
d_{0i}	4.47	3.61	8.06	9.49	6.71	8.94	13.54



- **Contouring par triangulation** : cette méthode consiste à établir l'équation du plan $z = a_0 + a_1x + a_2y$ formé par les trois points encadrant celui à estimer. Une **triangulation de Delaunay** montre que M_0 est encadré par les trois points M_2, M_3 et M_5 si bien qu'on peut établir le système de trois équations à trois inconnues suivant dont la solution unique est calculée par :

$$\begin{cases} a_0 + 63a_1 + 140a_2 = 696 \\ a_0 + 64a_1 + 129a_2 = 227 \\ a_0 + 71a_1 + 140a_2 = 606 \end{cases} \Rightarrow \begin{pmatrix} 1 & 63 & 140 \\ 1 & 64 & 129 \\ 1 & 71 & 140 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 696 \\ 227 \\ 606 \end{pmatrix} \Rightarrow \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -4421.16 \\ -11.25 \\ 41.61 \end{pmatrix}$$

On a donc $z = -4421.16 - 11.25x + 41.61y$ et $z_0 = 548.7$ ppm.

- **Interpolation polynômiale** : on cherche une fonction de la forme $f(x, y) = \sum_{j=0}^p a_j f_j(x, y)$ où les $f_j(x, y)$ sont des fonctions élémentaires de x et y : $1, x, y, xy, x^2, y^2, \dots$. Pour des interpolations locales, il est préconisé de se limiter à une fonction de la forme $f(x, y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2$ mais cette formulation peut être beaucoup plus complexe pour des analyses à l'échelle de la Terre par exemple. En posant cette équation pour les sept points de mesures on obtient le système matriciel suivant que l'on résout par la méthode des moindres carrés :

$$\begin{pmatrix} 1 & 61 & 139 & 61^2 & 61 \times 139 & 139^2 \\ 1 & 63 & 140 & 63^2 & 63 \times 140 & 140^2 \\ 1 & 64 & 129 & 64^2 & 64 \times 129 & 129^2 \\ 1 & 68 & 128 & 68^2 & 68 \times 128 & 128^2 \\ 1 & 71 & 140 & 71^2 & 71 \times 140 & 140^2 \\ 1 & 73 & 141 & 73^2 & 73 \times 141 & 141^2 \\ 1 & 75 & 128 & 75^2 & 75 \times 128 & 128^2 \end{pmatrix} \times \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} 477 \\ 696 \\ 227 \\ 646 \\ 606 \\ 791 \\ 783 \end{pmatrix} \Rightarrow \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} 312866.05 \\ 873.18 \\ -5123.24 \\ -4.57 \\ -1.85 \\ 20.29 \end{pmatrix}$$

On a donc $z_0 = 111.94$ ppm... ce qui n'est pas très réaliste. Conséquence : cette technique, bien adaptée à des champs dont les variations sont régulières, peut s'avérer médiocre selon les fonctions de base choisies.

IV.3. Ajustement à une fonction quelconque

Considérons une fonction mathématique f de p paramètres $\theta_1, \dots, \theta_p$. Jusqu'à présent, nous avons choisi des fonctions s'écrivant comme des combinaisons linéaires de $\theta_1, \dots, \theta_p$ si bien que les dérivées partielles du χ^2 par rapport à $\theta_1, \dots, \theta_p$ formaient un système d'équations dont la solution était immédiate. Si f est quelconque, cette approche n'est plus valable. Nous allons appliquer la **méthode de quasi-Newton** qui consiste à linéariser la fonction à minimiser. Pour cela, on écrit le développement en **série de Taylor** au **second ordre** de χ^2 en $\Theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_p^0)$:

$$\begin{aligned}\chi^2(\theta_1, \theta_2, \dots, \theta_p) &= \chi^2(\theta_1^0, \theta_2^0, \dots, \theta_p^0) + \sum_{j=1}^p (\theta_j - \theta_j^0) \frac{\partial \chi^2}{\partial \theta_j}(\theta_1^0, \theta_2^0, \dots, \theta_p^0) + \frac{1}{2!} \sum_{j=1}^p \sum_{k=1}^p (\theta_j - \theta_j^0)(\theta_k - \theta_k^0) \frac{\partial^2 \chi^2}{\partial \theta_j \partial \theta_k}(\theta_1^0, \theta_2^0, \dots, \theta_p^0) \\ &= \chi^2|_{\Theta^0} + \sum_{j=1}^p (\theta_j - \theta_j^0) \frac{\partial \chi^2}{\partial \theta_j} \Big|_{\Theta^0} + \frac{1}{2!} \sum_{j=1}^p \sum_{k=1}^p (\theta_j - \theta_j^0)(\theta_k - \theta_k^0) \frac{\partial^2 \chi^2}{\partial \theta_j \partial \theta_k} \Big|_{\Theta^0}\end{aligned}$$

En posant $\delta\theta_j = \theta_j - \theta_j^0$, χ^2 s'exprime directement linéairement en fonction des $\delta\theta_j$:

$$\chi^2(\theta_1, \theta_2, \dots, \theta_p) = \chi^2|_{\Theta^0} + \sum_{j=1}^p \delta\theta_j \frac{\partial \chi^2}{\partial \theta_j} \Big|_{\Theta^0} + \frac{1}{2} \sum_{j=1}^p \sum_{k=1}^p \delta\theta_j \delta\theta_k \frac{\partial^2 \chi^2}{\partial \theta_j \partial \theta_k} \Big|_{\Theta^0}$$

On applique alors la méthode des moindres carrés en annulant les dérivées partielles de χ^2 par rapport à chacun des paramètres. Prenons le cas de θ_1 :

$$\chi^2|_{\Theta^0} = \chi^2(\theta_1^0, \theta_2^0, \dots, \theta_p^0) \Rightarrow \frac{\partial}{\partial \theta_1} (\chi^2|_{\Theta^0}) = 0$$

$$\sum_{j=1}^p \delta\theta_j \frac{\partial \chi^2}{\partial \theta_j} \Big|_{\Theta^0} = (\theta_1 - \theta_1^0) \frac{\partial \chi^2}{\partial \theta_1} \Big|_{\Theta^0} + (\theta_2 - \theta_2^0) \frac{\partial \chi^2}{\partial \theta_2} \Big|_{\Theta^0} + \dots + (\theta_p - \theta_p^0) \frac{\partial \chi^2}{\partial \theta_p} \Big|_{\Theta^0} \Rightarrow \frac{\partial}{\partial \theta_1} \left(\sum_{j=1}^p \delta\theta_j \frac{\partial \chi^2}{\partial \theta_j} \Big|_{\Theta^0} \right) = \frac{\partial \chi^2}{\partial \theta_1} \Big|_{\Theta^0}$$

$$\begin{aligned}\sum_{j=1}^p \sum_{k=1}^p \delta\theta_j \delta\theta_k \frac{\partial^2 \chi^2}{\partial \theta_j \partial \theta_k} \Big|_{\Theta^0} &= (\theta_1 - \theta_1^0) \sum_{k=1}^p (\theta_k - \theta_k^0) \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_k} \Big|_{\Theta^0} + (\theta_2 - \theta_2^0) \sum_{k=1}^p (\theta_k - \theta_k^0) \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_k} \Big|_{\Theta^0} + \dots + (\theta_p - \theta_p^0) \sum_{k=1}^p (\theta_k - \theta_k^0) \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_k} \Big|_{\Theta^0} \\ \Rightarrow \frac{\partial}{\partial \theta_1} \left(\sum_{j=1}^p \sum_{k=1}^p \delta\theta_j \delta\theta_k \frac{\partial^2 \chi^2}{\partial \theta_j \partial \theta_k} \Big|_{\Theta^0} \right) &= 2 \left((\theta_1 - \theta_1^0) \frac{\partial^2 \chi^2}{\partial \theta_1^2} \Big|_{\Theta^0} + (\theta_2 - \theta_2^0) \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_2} \Big|_{\Theta^0} + \dots + (\theta_p - \theta_p^0) \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_p} \Big|_{\Theta^0} \right) = 2 \sum_{j=1}^p \delta\theta_j \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_j} \Big|_{\Theta^0}\end{aligned}$$

On fait de même avec $\theta_2, \theta_3, \dots, \theta_p$ et on obtient le système d'équations suivant :

$$\begin{cases} \frac{\partial \chi^2}{\partial \theta_1} = \frac{\partial \chi^2}{\partial \theta_1} \Big|_{\Theta^0} + \sum_{j=1}^p \delta\theta_j \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_j} \Big|_{\Theta^0} = 0 & \frac{\partial \chi^2}{\partial \theta_1} = \frac{\partial \chi^2}{\partial \theta_1} \Big|_{\Theta^0} \\ \frac{\partial \chi^2}{\partial \theta_2} = \frac{\partial \chi^2}{\partial \theta_2} \Big|_{\Theta^0} + \sum_{j=1}^p \delta\theta_j \frac{\partial^2 \chi^2}{\partial \theta_2 \partial \theta_j} \Big|_{\Theta^0} = 0 & \frac{\partial \chi^2}{\partial \theta_2} = \frac{\partial \chi^2}{\partial \theta_2} \Big|_{\Theta^0} \\ \dots & \dots \\ \frac{\partial \chi^2}{\partial \theta_p} = \frac{\partial \chi^2}{\partial \theta_p} \Big|_{\Theta^0} + \sum_{j=1}^p \delta\theta_j \frac{\partial^2 \chi^2}{\partial \theta_p \partial \theta_j} \Big|_{\Theta^0} = 0 & \frac{\partial \chi^2}{\partial \theta_p} = \frac{\partial \chi^2}{\partial \theta_p} \Big|_{\Theta^0} \end{cases} \Rightarrow \begin{cases} \sum_{j=1}^p \delta\theta_j \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_j} \Big|_{\Theta^0} = - \frac{\partial \chi^2}{\partial \theta_1} \Big|_{\Theta^0} \\ \sum_{j=1}^p \delta\theta_j \frac{\partial^2 \chi^2}{\partial \theta_2 \partial \theta_j} \Big|_{\Theta^0} = - \frac{\partial \chi^2}{\partial \theta_2} \Big|_{\Theta^0} \\ \dots \\ \sum_{j=1}^p \delta\theta_j \frac{\partial^2 \chi^2}{\partial \theta_p \partial \theta_j} \Big|_{\Theta^0} = - \frac{\partial \chi^2}{\partial \theta_p} \Big|_{\Theta^0} \end{cases}$$

On peut écrire ce système sous forme matricielle :

$$\begin{pmatrix} \frac{\partial^2 \chi^2}{\partial \theta_1^2} \Big|_{\Theta^0} & \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_2} \Big|_{\Theta^0} & \dots & \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_p} \Big|_{\Theta^0} \\ \frac{\partial^2 \chi^2}{\partial \theta_2 \partial \theta_1} \Big|_{\Theta^0} & \frac{\partial^2 \chi^2}{\partial \theta_2^2} \Big|_{\Theta^0} & \dots & \frac{\partial^2 \chi^2}{\partial \theta_2 \partial \theta_p} \Big|_{\Theta^0} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \chi^2}{\partial \theta_p \partial \theta_1} \Big|_{\Theta^0} & \frac{\partial^2 \chi^2}{\partial \theta_p \partial \theta_2} \Big|_{\Theta^0} & \dots & \frac{\partial^2 \chi^2}{\partial \theta_p^2} \Big|_{\Theta^0} \end{pmatrix} \times \begin{pmatrix} \delta\theta_1 \\ \delta\theta_2 \\ \vdots \\ \delta\theta_p \end{pmatrix} = - \begin{pmatrix} \frac{\partial \chi^2}{\partial \theta_1} \Big|_{\Theta^0} \\ \frac{\partial \chi^2}{\partial \theta_2} \Big|_{\Theta^0} \\ \vdots \\ \frac{\partial \chi^2}{\partial \theta_p} \Big|_{\Theta^0} \end{pmatrix}$$

et donc la solution $\Theta = (\theta_1, \theta_2, \dots, \theta_k)$ s'écrit :

$$\begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{pmatrix} = \begin{pmatrix} \theta_1^0 \\ \theta_2^0 \\ \vdots \\ \theta_p^0 \end{pmatrix} - \left(\begin{array}{ccc} \frac{\partial^2 \chi^2}{\partial \theta_1^2} \Big|_{\Theta^0} & \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_2} \Big|_{\Theta^0} & \dots & \frac{\partial^2 \chi^2}{\partial \theta_1 \partial \theta_p} \Big|_{\Theta^0} \\ \frac{\partial^2 \chi^2}{\partial \theta_2 \partial \theta_1} \Big|_{\Theta^0} & \frac{\partial^2 \chi^2}{\partial \theta_2^2} \Big|_{\Theta^0} & \dots & \frac{\partial^2 \chi^2}{\partial \theta_2 \partial \theta_p} \Big|_{\Theta^0} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \chi^2}{\partial \theta_p \partial \theta_1} \Big|_{\Theta^0} & \frac{\partial^2 \chi^2}{\partial \theta_p \partial \theta_2} \Big|_{\Theta^0} & \dots & \frac{\partial^2 \chi^2}{\partial \theta_p^2} \Big|_{\Theta^0} \end{array} \right)^{-1} \times \begin{pmatrix} \frac{\partial \chi^2}{\partial \theta_1} \Big|_{\Theta^0} \\ \frac{\partial \chi^2}{\partial \theta_2} \Big|_{\Theta^0} \\ \vdots \\ \frac{\partial \chi^2}{\partial \theta_k} \Big|_{\Theta^0} \end{pmatrix} \Rightarrow \Theta = \Theta^0 - H^{-1} \Big|_{\Theta^0} \times \nabla \chi^2 \Big|_{\Theta^0}$$

La matrice des dérivées partielles notée H est appelée **matrice Hessienne** : c'est une matrice carrée et symétrique d'ordre p. Le vecteur des dérivées noté $\nabla \chi^2$ est le **vecteur gradient** de dimension p×1. En général un seul calcul n'est pas suffisant pour trouver la solution et il faut procéder par itérations successives. On part d'un jeu de variables initial $\Theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_p^0)$ et on calcule $\Theta^1 = (\theta_1^1, \theta_2^1, \dots, \theta_p^1)$, etc. jusqu'à ce qu'un test de convergence soit vérifié.

Références

- Arbenz K., Wohlhauser A. (1996), *Analyse numérique*, Presses Polytechniques et Universitaires Romandes, 84 pages.
- Bevington P.R., Robinson D.K. (1992), *Data reduction and error analysis for the physical sciences*, McGraw-Hill, 328 pages.
- Borse G.J. (1997), *Numerical methods with MATLAB: A resource for scientists and engineers*, PWS Publishing Company, 640 pages.
- Huré J.-M., Pelat D. (2002), *Méthodes numériques : éléments d'un premier parcours*, Cours du DEA « Astrophysique & Méthodes Associées », 110 pages.
- Isaaks E.H., Srivastava R.M. (1989), *An introduction to applied geostatistics*, Oxford University Press, 561 pages.
- Middleton G.V. (2000), *Data analysis in the Earth sciences using MATLAB*, Prentice-Hall, 260 pages.
- Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P. (1992), *Numerical recipes in Fortran - The art of scientific computing*, Cambridge University Press, 963 pages.
- Quarteroni A. (2000), *Méthodes numériques pour le calcul scientifique – Programmes en Matlab*, Springer, 444 pages.
- Viot P. (2003), *Méthodes d'analyse numérique*, Cours du DEA « Modélisation Dynamique et Statistique des Systèmes Complexes », 108 pages.