Chapitre II

Formes quadratiques.

Polynômes vectoriels du 2nd degré.

1. Formes quadratiques
   - Bases orthonormales et matrices orthogonales
   - Projection orthogonale
   - Procédé d'orthonormalisation de Gram-Schmidt
   - Formes bilinéaires
   - Formes quadratiques
   - Théorème spectral
   - Matrices symétriques définies positives
   - Application : optimisation sous contrainte – quotient de Rayleigh

2. Polynômes vectoriels du 2nd degré
   - Définitions
   - Gradient
   - Optimisation

1. Formes quadratiques

Bases orthonormales et matrices orthogonales

- $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$
- E est un espace euclidien c'est-à-dire un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ et d'un produit scalaire $\langle . , . \rangle$

Une base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ est **orthonormale** si $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij}$.

$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ *symbole de Kronecker*

Dans ce cas pour $\mathbf{u}$ et $\mathbf{v}$ dans E :

$$\mathbf{u} = \sum_{i=1}^{n} u_i \mathbf{e}_i$$

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} u_i v_i$$

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\sum_{i=1}^{n} u_i^2}$$

***norme de* $\mathbf{u}$**

avec $u_i = \langle \mathbf{u}, \mathbf{e}_i \rangle$ et $v_i = \langle \mathbf{v}, \mathbf{e}_i \rangle$

On rappelle que :

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| . \|\mathbf{v}\|$$

Inégalité de Schwarz

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Inégalité triangulaire

1. Formes quadratiques

Bases orthonormales et matrices orthogonales

Exemples

- $E = \mathbb{R}^n$ muni de la base canonique $\mathcal{B} = \{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_n\}$ et du produit scalaire défini par

$$\boxed{\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} x_i y_i}$$

avec

$$\mathbf{x} = \sum_{i=1}^{n} x_i \boldsymbol{\varepsilon}_i = (x_1, x_2, \dots, x_n)^T \quad \text{et} \quad \mathbf{y} = \sum_{i=1}^{n} y_i \boldsymbol{\varepsilon}_i = (y_1, y_2, \dots, y_n)^T$$

- $E = \mathbb{R}^3$, $\mathbf{e}_1 = \frac{\sqrt{3}}{3}(1,1,1)^T$, $\mathbf{e}_2 = \frac{\sqrt{6}}{6}(-2,1,1)^T$, $\mathbf{e}_3 = \frac{\sqrt{2}}{2}(0,-1,1)^T$

$\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ est une base orthonormale.

1. Formes quadratiques

Bases orthonormales et matrices orthogonales

On s'intéresse à la matrice de passage $\mathbf{P}$ entre deux bases orthonormales $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ et $\mathcal{B}' = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$.

$$\mathbf{P} = \left(p_{ij}\right) = \begin{array}{c} \\ \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{f}_1 & \mathbf{f}_2 & \dots & \mathbf{f}_n \end{array} \\ \left(\begin{array}{c|c|c|c} & & & \\ & & & \\ & & & \\ & & & \end{array} \right) \end{array}$$

$$\left.\begin{aligned} &\text{Donc : } \mathbf{f}_k = \mathbf{P}\mathbf{e}_k = \sum_{i=1}^{n} p_{ik}\mathbf{e}_i \\ &\text{Ainsi : } \langle \mathbf{f}_k, \mathbf{e}_j \rangle = \left\langle \sum_{i=1}^{n} p_{ik}\mathbf{e}_i \, , \mathbf{e}_j \right\rangle = \sum_{i=1}^{n} p_{ik} \underbrace{\langle \mathbf{e}_i, \mathbf{e}_j \rangle}_{\delta_{ij}} = p_{jk} \end{aligned}\right\} \quad \mathbf{P} = \left(p_{ij}\right) \quad \text{avec} \quad p_{ij} = \langle \mathbf{f}_j, \mathbf{e}_i \rangle$$

On considère maintenant la matrice transposée de $\mathbf{P}$, c'est-à-dire $\mathbf{P}^T = \left(q_{ij}\right)$ avec $q_{ij} = p_{ji}$:

$$\left.\begin{aligned} &\text{Donc : } \quad \mathbf{P}^T\mathbf{f}_k = \sum_{i=1}^{n} q_{ik}\mathbf{f}_i = \sum_{i=1}^{n} p_{ki}\mathbf{f}_i = \sum_{i=1}^{n} \langle \mathbf{f}_i \, , \mathbf{e}_k \rangle \mathbf{f}_i = \mathbf{e}_k \\ &\text{Par ailleurs : } \quad \mathbf{P}^T\mathbf{f}_k = \mathbf{P}^T\mathbf{P}\mathbf{e}_k \end{aligned}\right\} \quad \mathbf{P}^T\mathbf{P} = \mathbf{I}$$

1. Formes quadratiques

Bases orthonormales et matrices orthogonales

Définition : Les matrices vérifiant $\mathbf{P}^T\mathbf{P} = \mathbf{I}$ sont appelées **matrices orthogonales**.

Remarques

- $\mathbf{P}$ est inversible (puisque c'est une matrice de passage) donc $\mathbf{P}^{-1} = \mathbf{P}^T$.

- De plus : $\det(\mathbf{P}^T\mathbf{P}) = 1 \;\Rightarrow\; \det\mathbf{P}^T \det\mathbf{P} = 1 \;\Rightarrow\; (\det\mathbf{P})^2 = 1 \;\Rightarrow\; \det\mathbf{P} = \pm 1$

  Attention la réciproque est fausse : $\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1$ et $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ n'est pas orthogonale.

- $\langle \mathbf{Px}, \mathbf{Py} \rangle = (\mathbf{Px})^T\mathbf{Py} = \mathbf{x}^T\mathbf{P}^T\mathbf{Py} = \mathbf{x}^T\mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$

  En en particulier, si $\mathbf{y} = \mathbf{x}$: $\|\mathbf{Px}\| = \|\mathbf{x}\|$

  $\Rightarrow$ une matrice orthogonale <u>conserve le produit scalaire</u>, on dit que c'est une **isométrie**.

1. Formes quadratiques

Bases orthonormales et matrices orthogonales

En résumé

- *Une matrice orthogonale est une matrice qui a pour inverse sa transposée.*
- *La matrice de passage d'une base orthonormale à une base orthonormale est une matrice orthogonale.*
- *Réciproquement, toute matrice orthogonale transforme une base orthonormale en une base orthonormale.*
- *Si les vecteurs colonnes d'une matrice sont orthonormaux alors la matrice est orthogonale et réciproquement (même propriété pour les lignes).*
- *Le déterminant d'une matrice orthogonale est égal à* ± 1.
- *Une matrice orthogonale est une isométrie.*

1. Formes quadratiques

Bases orthonormales et matrices orthogonales

Soit $\mathbf{A}$ une matrice <u>symétrique</u> dans une base orthonormale et $\mathbf{Q}$ la matrice de passage à une autre base orthonormale, alors la matrice semblable à $\mathbf{A}$ dans la nouvelle base s'écrit (cf. chapitre 1) :

$$\mathbf{A}' = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{Q}^{T}\mathbf{A}\mathbf{Q}$$

Donc :

$$(\mathbf{A}')^{T} = (\mathbf{Q}^{T}\mathbf{A}\mathbf{Q})^{T} = \mathbf{Q}^{T}\mathbf{A}^{T}\mathbf{Q} = \mathbf{Q}^{T}\mathbf{A}\mathbf{Q} = \mathbf{A}'$$

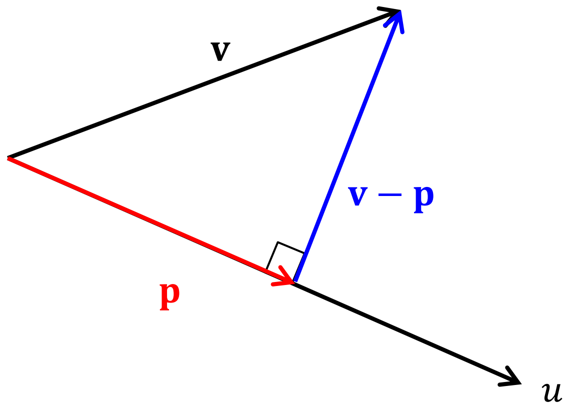
La matrice $\mathbf{A}'$ est également symétrique.

A retenir

Tout changement de base par une matrice orthogonale transforme une matrice symétrique en une matrice symétrique.

1. Formes quadratiques

Projection orthogonale



On considère 2 vecteurs $\mathbf{u}$ et $\mathbf{v}$ de E et on cherche le vecteur $\mathbf{p}$ tel que $\mathbf{p}$ soit la projection orthogonale de $\mathbf{v}$ sur $\mathbf{u}$.

Pour cela, il faut que :

$$\mathbf{p} = \alpha\mathbf{u}, \quad \alpha \in \mathbb{R} \qquad \text{et} \qquad \langle \mathbf{v}-\mathbf{p}, \mathbf{p}\rangle = 0$$

On a alors : $\langle \mathbf{v}-\mathbf{p}, \mathbf{p}\rangle = 0 \quad \Leftrightarrow \quad \langle \mathbf{v}, \mathbf{p}\rangle - \langle \mathbf{p}, \mathbf{p}\rangle = 0$

$$\Leftrightarrow \quad \alpha\langle \mathbf{v}, \mathbf{u}\rangle - \alpha^2\langle \mathbf{u}, \mathbf{u}\rangle = 0 \quad \Leftrightarrow \quad \alpha = \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\|\mathbf{u}\|^2}$$

A retenir

La projection orthogonale d'un vecteur $\mathbf{v}$ *sur un vecteur* $\mathbf{u}$ *est égale à :*

$$\mathbf{p} = \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\|\mathbf{u}\|^2}\mathbf{u}$$

1. Formes quadratiques

Procédé d'orthonormalisation de Gram-Schmidt

Exemple

On considère les 3 vecteurs $\mathbf{e}_1 = (1,1,1)^T$, $\mathbf{e}_2 = (0,1,1)^T$, $\mathbf{e}_3 = (0,0,1)^T$

On vérifie facilement que $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ forme une famille libre dans $\mathbb{R}^3$ et par conséquent une base de $\mathbb{R}^3$.

La base $\mathcal{B}$ n'est pas orthogonale (les vecteurs $\mathbf{e}_i$ ne sont ni orthogonaux deux à deux, ni de norme 1).

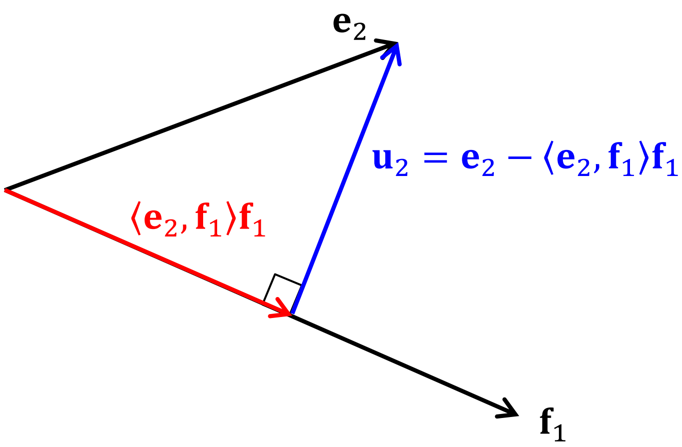
On cherche un procédé qui permet de transformer la base $\mathcal{B}$ en une base $\mathcal{B}' = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ orthonormale.

<u>1$^{\text{ère}}$ étape</u> :

$$\mathbf{f}_1 = \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|} = \frac{(1,1,1)^T}{\sqrt{3}} = \frac{\sqrt{3}}{3}(1,1,1)^T \qquad \Rightarrow \quad \|\mathbf{f}_1\| = 1$$

Procédé d'orthonormalisation de Gram-Schmidt

2ème étape :



$$\mathbf{u}_2 = \mathbf{e}_2 - \langle \mathbf{e}_2, \mathbf{f}_1 \rangle \mathbf{f}_1$$

$$\mathbf{u}_2 = (0,1,1)^T - \frac{2\sqrt{3}}{3} \times \frac{\sqrt{3}}{3} (1,1,1)^T = \frac{1}{3} (-2,1,1)^T \qquad \Rightarrow \quad \langle \mathbf{u}_2, \mathbf{f}_1 \rangle = 0$$

$$\mathbf{f}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{\frac{1}{3}(-2,1,1)^T}{\sqrt{\frac{2}{3}}} = \frac{\sqrt{6}}{6} (-2,1,1)^T \qquad \Rightarrow \quad \langle \mathbf{f}_2, \mathbf{f}_1 \rangle = 0 \quad \text{et} \quad \|\mathbf{f}_2\| = 1$$

1. Formes quadratiques

Procédé d'orthonormalisation de Gram-Schmidt

3ème étape :

$$\mathbf{u}_3 = \mathbf{e}_3 - \langle \mathbf{e}_3, \mathbf{f}_1 \rangle \mathbf{f}_1 - \langle \mathbf{e}_3, \mathbf{f}_2 \rangle \mathbf{f}_2$$

$$\mathbf{u}_3 = (0,0,1)^T - \frac{\sqrt{3}}{3} \times \frac{\sqrt{3}}{3}(1,1,1)^T - \frac{\sqrt{6}}{6} \times \frac{\sqrt{6}}{6}(-2,1,1)^T = \cdots = \frac{1}{2}(0,-1,1)^T$$

$$\Rightarrow \quad \langle \mathbf{u}_3, \mathbf{f}_1 \rangle = 0 \qquad \text{et} \qquad \langle \mathbf{u}_3, \mathbf{f}_2 \rangle = 0$$

$$\mathbf{f}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = \frac{\frac{1}{2}(0,-1,1)^T}{\frac{\sqrt{2}}{2}} = \frac{\sqrt{2}}{2}(0,-1,1)^T$$

$$\Rightarrow \quad \langle \mathbf{f}_3, \mathbf{f}_1 \rangle = 0, \qquad \langle \mathbf{f}_3, \mathbf{f}_2 \rangle = 0 \qquad \text{et} \qquad \|\mathbf{f}_3\| = 1$$

Conclusion : $\mathcal{B}' = \left\{ \frac{\sqrt{3}}{3}(1,1,1)^T,\ \frac{\sqrt{6}}{6}(-2,1,1)^T,\ \frac{\sqrt{2}}{2}(0,-1,1)^T \right\}$ est une base orthonormale de $\mathbb{R}^3$.

1. Formes quadratiques

Procédé d'orthonormalisation de Gram-Schmidt

L'exemple précédent est une illustration du **procédé d'orthonormalisation de Gram-Schmidt** que l'on peut énoncé de la façon suivante :

Soit $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_\mathrm{n}\}$ une base de E alors il existe une unique base orthonormale $\mathcal{B}' = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_\mathrm{n}\}$ telle que :

$$\forall k \in [\![1, n]\!], \qquad \mathrm{Vect}(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k) = \mathrm{Vect}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$$

Les vecteurs $\mathbf{f}_\mathrm{i}$ de la nouvelle base sont définis par :

$$\mathbf{f}_1 = \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|} \qquad \text{et} \qquad \forall i \in [\![2, n]\!], \qquad \mathbf{f}_i = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|}$$

$$\text{avec} \quad \mathbf{u}_i = \mathbf{e}_i - \sum_{j=1}^{i-1} \langle \mathbf{e}_i, \mathbf{f}_j \rangle \mathbf{f}_j$$

1. Formes quadratiques

Forme bilinéaire

- $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$
- E désigne un $\mathbb{K}$-espace vectoriel de dimension n muni d'une base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$

Une **forme bilinéaire** sur E est une application

$$\varphi\colon E \times E \longrightarrow \mathbb{K}$$
$$(\mathbf{x}, \mathbf{y}) \longmapsto \varphi(\mathbf{x}, \mathbf{y})$$

telle que :

$$\varphi(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2, \mathbf{y}) = \alpha\varphi(\mathbf{x}_1, \mathbf{y}) + \beta\varphi(\mathbf{x}_2, \mathbf{y})$$

$$\varphi(\mathbf{x}, \alpha\mathbf{y}_1 + \beta\mathbf{y}_2) = \alpha\varphi(\mathbf{x}, \mathbf{y}_1) + \beta\varphi(\mathbf{x}, \mathbf{y}_2)$$

Exemple : $E = \mathbb{R}^{\boldsymbol{n}}$, $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in E$ et $\mathbf{y} = (y_1, y_2, \dots, y_n)^T \in E$

$$\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y} = \sum_{k=1}^{n} x_k y_k$$ ***produit scalaire***

1. Formes quadratiques

Forme bilinéaire

Matrice associée à une forme bilinéaire

$\mathbf{x}$ et $\mathbf{y}$ peuvent aussi s'écrire : $\quad \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \quad$ et $\quad \mathbf{y} = \sum_{j=1}^{n} y_j \mathbf{e}_j$

Ainsi :

$$\varphi(\mathbf{x}, \mathbf{y}) = \varphi\left(\sum_{i=1}^{n} x_i \mathbf{e}_i, \sum_{j=1}^{n} y_j \mathbf{e}_j\right) = \sum_{i=1}^{n} x_i \varphi\left(\mathbf{e}_i, \sum_{j=1}^{n} y_j \mathbf{e}_j\right) = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} y_j \varphi(\mathbf{e}_i, \mathbf{e})\right)$$

$$= \sum_{i,j} x_i \varphi(\mathbf{e}_i, \mathbf{e}_j) y_j$$

$$= \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} \varphi(\mathbf{e}_1, \mathbf{e}_1) & \varphi(\mathbf{e}_1, \mathbf{e}_2) & \dots & \varphi(\mathbf{e}_1, \mathbf{e}_n) \\ \varphi(\mathbf{e}_2, \mathbf{e}_1) & \varphi(\mathbf{e}_2, \mathbf{e}_2) & \dots & \varphi(\mathbf{e}_2, \mathbf{e}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(\mathbf{e}_n, \mathbf{e}_1) & \varphi(\mathbf{e}_n, \mathbf{e}_2) & \dots & \varphi(\mathbf{e}_n, \mathbf{e}_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \mathbf{x}^T \, \mathbf{A}\mathbf{y} \qquad \text{avec} \qquad \mathbf{A} = \left(\varphi(\mathbf{e}_i, \mathbf{e}_j)\right)_{\substack{1 \le i \le n \\ 1 \le j \le n}}$$

1. Formes quadratiques

Forme bilinéaire

Remarques

- $\varphi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A}\mathbf{y}$ *est le produit scalaire des vecteurs* $\mathbf{x}$ *et* $\mathbf{A}\mathbf{y}$

- $\varphi(\mathbf{x}, \mathbf{y})$ *est un scalaire (réel ou complexe) donc* $\varphi(\mathbf{x}, \mathbf{y}) = \left(\varphi(\mathbf{x}, \mathbf{y})\right)^T = (\mathbf{x}^T \mathbf{A}\mathbf{y})^T = \mathbf{y}^T \mathbf{A}^T \mathbf{x}$ *donc*

  $\varphi(\mathbf{x}, \mathbf{y})$ *est aussi le produit scalaire de* $\mathbf{x}$ *et* $\mathbf{A}^T \mathbf{y}$

1. Formes quadratiques

Forme bilinéaire

Exemple : $\varphi(\mathbf{x},\mathbf{y}) = x_1y_1 - x_1y_2 + 2x_2y_1 + 3x_2y_2$

ou

1. Formes quadratiques

Forme quadratique

Une forme bilinéaire **symétrique** φ est définie par : $\varphi(\mathbf{x},\mathbf{y}) = \varphi(\mathbf{y},\mathbf{x})$

$$\left.\begin{array}{l} \bullet\ \ \varphi(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y} \\ \\ \bullet\ \ \varphi(\mathbf{y},\mathbf{x}) = \mathbf{y}^T\mathbf{A}\mathbf{x} = (\mathbf{x}^T\mathbf{A}^T\mathbf{y})^T = \mathbf{x}^T\mathbf{A}^T\mathbf{y} \end{array}\right\} \Rightarrow \mathbf{A} = \mathbf{A}^T$$

La matrice associée à une forme bilinéaire symétrique est symétrique.

Une **forme quadratique** q associée à une forme bilinéaire symétrique φ est définie par :

$$\begin{aligned} q:\ E &\longrightarrow \mathbb{K} \\ \mathbf{x} &\longmapsto q(\mathbf{x}) = \varphi(\mathbf{x},\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} \end{aligned}$$

1. Formes quadratiques

Forme quadratique

Remarque

Une forme quadratique peut être données sous la forme : $q(\mathbf{x}) = \sum_{j=1}^{n} \sum_{i \leq j} \alpha_{ij} x_i x_j$

En effet :

$$q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j = \sum_{j=1}^{n} \sum_{i \leq j} a_{ij} x_i x_j + \sum_{j=1}^{n} \sum_{i > j} a_{ij} x_i x_j$$

$$= \sum_{j=1}^{n} \sum_{i \leq j} a_{ij} x_i x_j + \sum_{j=1}^{n} \sum_{i < j} a_{ij} x_i x_j$$

$$= 2 \sum_{j=1}^{n} \sum_{i < j} a_{ij} x_i x_j + \sum_{j=1}^{n} a_{jj} x_j^2$$

$$= \sum_{j=1}^{n} \sum_{i \leq j} \alpha_{ij} x_i x_j \qquad \text{où } \alpha_{ij} = \begin{cases} 2a_{ij}, & i < j \\ a_{jj}, & i = j \end{cases}$$

1. Formes quadratiques

Forme quadratique

Exemple : $q(\mathbf{x}) = x_1^2 + 2x_1x_3 - 2x_2x_3 + 2x_2^2 - x_3^2$

1. Formes quadratiques

Théorème spectral

Le théorème spectral, encore appelé **théorème d'orthodiagonalisation**, dit que toute matrice symétrique réelle $\mathbf{A}$ peut être orthodiagonalisée, c'est-à-dire qu'il existe une matrice orthogonale $\mathbf{Q}$ et une matrice diagonale $\mathbf{D}$ telles que:

$$\mathbf{A} = \mathbf{QDQ}^T$$

Remarques importantes

- *Les vecteurs propres de $\mathbf{A}$ forment une base orthonormale de $\mathbb{R}^n$ (car $\mathbf{Q}$ est orthogonale et les vecteurs propres de $\mathbf{A}$ sont les vecteurs colonnes de $\mathbf{Q}$)*

- *Si λ est une valeur propre de $\mathbf{A}$ associée au vecteur propre $\mathbf{x} \in \mathbb{R}^{\boldsymbol{n}}$ alors :*

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\lambda\mathbf{x} = \lambda\mathbf{x}^T\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

or $\mathbf{x}^T\mathbf{A}\mathbf{x} = q(\mathbf{x})$ est la forme quadratique associée à $\mathbf{A}$ et $\mathbf{x} \in \mathbb{R}^n$ donc $q(\mathbf{x}) \in \mathbb{R}$; comme par ailleurs $\|\mathbf{x}\|^2 \in \mathbb{R}$, on en déduit que $\lambda \in \mathbb{R}$.

1. Formes quadratiques

Théorème spectral

Remarques importantes (suite)

- *Considérons 2 valeurs propres distinctes λ et μ associées aux vecteurs propres $\mathbf{x}$ et $\mathbf{y}$ respectivement et si calculons $(\mathbf{Ax})^T\mathbf{y}$ de 2 façons différentes :*

  - $(\mathbf{Ax})^T\mathbf{y} = \mathbf{x}^T\mathbf{Ay} = \mu\mathbf{x}^T\mathbf{y}$
  - $(\mathbf{Ax})^T\mathbf{y} = (\lambda\mathbf{x})^T\mathbf{y} = \lambda\mathbf{x}^T\mathbf{y}$

  On en déduit que $(\mu-\lambda)\mathbf{x}^T\mathbf{y} = 0$ c'est-à-dire $\mathbf{x}^T\mathbf{y} = 0$ donc $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux.

- $q(\mathbf{x}) = \mathbf{x}^T\mathbf{Ax} = \mathbf{x}^T\mathbf{QDQ}^T\mathbf{x} = (\mathbf{Q}^T\mathbf{x})^T\mathbf{D}(\mathbf{Q}^T\mathbf{x}) = \mathbf{y}^T\mathbf{Dy}$

$$= \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{k=1}^{n} \lambda_i y_i^2$$

Ce dernier résultat est connu sous le nom de ***théorème des axes principaux***

1. Formes quadratiques

Théorème spectral

En résumé

Soit $\mathbf{A}$ *une matrice* *symétrique réelle* *et* q *la forme quadratique associée*

- $\mathbf{A}$ *peut s'écrire sous la forme* $\mathbf{QDQ}^T$ *où* $\mathbf{Q}$ *est une matrice orthogonale et* $\mathbf{D}$ *est une matrice diagonale.*

- *Les vecteurs propres de* $\mathbf{A}$ *forment une base orthonormale de* $\mathbb{R}^n$.

- *Les valeurs propres de* $\mathbf{A}$ *sont réelles*

- *Deux vecteurs propres de* $\mathbf{A}$ *associés à deux valeurs propres distinctes sont orthogonaux.*

- $$q(\mathbf{x}) = \tilde{q}(\mathbf{y}) = \sum_{k=1}^{n} \lambda_i y_i^2 \quad \text{avec} \quad \mathbf{y} = \mathbf{Q}^T\mathbf{x}$$

1. Formes quadratiques

Théorème spectral

En résumé (suite)

<u>Réduire en base orthonormée</u> une forme quadratique q signifie :

- *écrire $q(\mathbf{x})$ sous la forme $\mathbf{x}^T\mathbf{A}\mathbf{x}$ où $\mathbf{A}$ est symétrique*
- *Diagonaliser $\mathbf{A}$ (c'est-à-dire trouver les matrices $\mathbf{Q}$ et $\mathbf{D}$ telles que $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$ où $\mathbf{Q}$ est une matrice orthogonale et $\mathbf{D}$ est une matrice diagonale)*
- *écrire q sous la forme $q = \sum_{k=1}^{n} \lambda_i y_i^2$ où les y_i sont les composantes du vecteur $\mathbf{y} = \mathbf{Q}^T\mathbf{x}$ et où les λ_i sont les valeurs propres de $\mathbf{A}$.*

1. Formes quadratiques

Théorème spectral

- $E = \mathbb{R}^{\mathbf{2}}$
- $\mathbf{x} = (x_1, x_2\,)^T$

Exemple : $q(\mathbf{x}) = 3x_1^2 + 10x_1x_2 + 3x_2^2$

$$q(\mathbf{x}) = \ 3x_1^2 + 5x_1x_2 + 5x_2x_1 + 3x_2^2 \ \ = \ x_1(3x_1 + 5x_2) + x_2(5x_1 + 3x_2)$$

$$= \ (x_1 \quad x_2)\begin{pmatrix} 3x_1 + 5x_2 \\ 5x_1 + 3x_2 \end{pmatrix} \ = \ (x_1 \quad x_2)\begin{pmatrix} 3 & 5 \\ 5 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \ = \ \mathbf{x}^T\mathbf{A}\mathbf{x} \qquad \text{avec } \mathbf{A} = \begin{pmatrix} 3 & 5 \\ 5 & 3 \end{pmatrix}$$

- valeurs propres de $\mathbf{A}$: $\lambda_1 = -2$ et $\lambda_2 = 8$
- vecteurs propres associés : $\mathbf{x}_1 = (1, -1)^T$ et $\mathbf{x}_2 = (1,1)^T$
- $\mathbf{x}_1$ et $\mathbf{x}_2$ sont orthogonaux (on a bien $\mathbf{x}_1^T\mathbf{x}_2 = 0$)
- on les normalise (en les divisant par leur norme) et on obtient la matrice orthogonale :

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

1. Formes quadratiques

Théorème spectral

- on pose $\mathbf{y} = \mathbf{Q}^T\mathbf{x} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} x_1 - x_2 \\ x_1 + x_2 \end{pmatrix}$

- finalement, d'après le théorème des axes principaux, on a :

$$q(\mathbf{x}) = \tilde{q}(\mathbf{y}) = -2y_1^2 + 8y_2^2$$

avec $y_1 = \frac{1}{\sqrt{2}}(x_1 - x_2)$ et $y_2 = \frac{1}{\sqrt{2}}(x_1 + x_2)$

1. Formes quadratiques

Matrice symétrique définie positive

Une matrice $\mathbf{A}$ est dite **symétrique définie positive** si la forme quadratique associée est positive c'est-à-dire si :

$$\forall \mathbf{x} \neq \mathbf{0}_E \ , \qquad \mathbf{x}^T \mathbf{A} \mathbf{x} > 0$$

Remarques importantes

- *Si $\mathbf{x}$ est un vecteur propre de $\mathbf{A}$ associé à la valeur propre λ alors :*

$$0 < \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \lambda \mathbf{x} = \lambda \|\mathbf{x}\|^2$$

Les valeurs propres d'un matrice symétrique définie positive sont donc <u>strictement positives</u>. La réciproque est vraie : toute matrice symétrique dont les valeurs propres sont strictement positives est définie positive.

On en déduit aussi que le déterminant d'une matrice symétrique définie positive est <u>strictement positif</u> puisque le déterminant est le produit des valeurs propres.

1. Formes quadratiques

Matrice symétrique définie positive

Remarques importantes (suite)

- *Supposons que* $\mathbf{A}$ *ne soit pas inversible alors :*

$$\mathrm{Ker}\,\mathbf{A} \neq \{\mathbf{0}\} \quad \Rightarrow \quad \exists \mathbf{x} \neq \mathbf{0}, \quad \mathbf{A}\mathbf{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{x}^T\mathbf{A}\mathbf{x} = 0$$

Une matrice symétrique définie positive est donc <u>inversible</u>.

- *On note* λ_i *ses valeurs propres de* $\mathbf{A}$. *On sait que* $\lambda_i > 0$ *et que* $\mathbf{A}$ *est inversible, donc* $\mathbf{A}^{-1}$ *existe et ses valeurs propres sont* $1/\lambda_i$ *: elles sont donc strictement positives. On en déduit que* $\mathbf{A}^{-1}$ *est aussi une matrice symétrique définie positive.*

1. Formes quadratiques

Matrice symétrique définie positive

Remarques importantes (suite et fin)

- *On note* $\mathbf{A}_i$ *les sous-matrices principales de* $\mathbf{A}$*, c'est-à-dire les blocs carrés supérieurs gauches.*

  Par exemple :

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} \Rightarrow \quad \mathbf{A}_1 = (2), \quad \mathbf{A}_2 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \quad \mathbf{A}_4 = \mathbf{A}$$

  Alors les matrices $\mathbf{A}_i$ *sont symétriques (évident) et* $\det \mathbf{A}_i > 0$

- *Pour finir :* $\mathbf{A} = \mathbf{QDQ}^T = (\mathbf{Q\Lambda})(\mathbf{\Lambda Q}^T) = (\mathbf{\Lambda Q}^T)^T(\mathbf{\Lambda Q}^T) = \mathbf{M}^T\mathbf{M}$

Donc toute matrice symétrique définie positive s'écrit $\mathbf{A} = \mathbf{M}^T\mathbf{M}$ *où* $\mathbf{M}$ *est une matrice inversible.*

1. Formes quadratiques

Matrice symétrique définie positive

A retenir

Une matrice $\mathbf{A}$ *est dite symétrique définie positive si elle vérifie l'une des propriétés équivalentes suivantes :*

- $\forall \mathbf{x} \neq \mathbf{0}_E \ , \ \mathbf{x}^T \mathbf{A} \mathbf{x} > 0$
- *Les valeurs propres de* $\mathbf{A}$ *sont réelles strictement positives*
- $\mathbf{A}$ *est inversible et* $\mathbf{A}^{-1}$ *est une matrice symétrique définie positive*
- $\mathbf{A} = \mathbf{M}^T \mathbf{M}$ *où* $\mathbf{M}$ *est une matrice inversible*
- *Toutes les sous-matrices principales de* $\mathbf{A}$ *sont symétriques définies positives*

On retiendra aussi que le déterminant d'une matrice symétrique définie positive est strictement positif (mais que la réciproque n'est pas nécessairement vraie).

1. Formes quadratiques

Application : optimisation sous contrainte – quotient de Rayleigh

Objectif : optimiser (minimiser ou maximiser) une forme quadratique $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x},\ \ \mathbf{x} \in \mathbb{R}^n$, sous la contrainte $g(\mathbf{x}) = \mathbf{x}^T\mathbf{x} - \alpha^2 = 0$ (hypersphère de rayon α).

Cela revient à optimiser le **quotient de Rayleigh** défini par :

$$\boxed{R(\mathbf{x}) = \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}}$$

D'après le théorème spectral :

$$\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$$

avec
- $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ ($\mathbf{Q}$ est orthogonale)
- $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ (λ_i : valeurs propres de $\mathbf{A}$)

Ainsi : $R(\mathbf{x}) = \dfrac{\mathbf{x}^T\mathbf{Q}\mathbf{D}\mathbf{Q}^T\mathbf{x}}{\mathbf{x}^T\mathbf{Q}\mathbf{Q}^T\mathbf{x}} = \dfrac{(\mathbf{Q}^T\mathbf{x})^T\mathbf{D}(\mathbf{Q}^T\mathbf{x})}{(\mathbf{Q}^T\mathbf{x})^{\boldsymbol{T}}(\mathbf{Q}^T\mathbf{x})} = \dfrac{\mathbf{y}^T\mathbf{D}\mathbf{y}}{\mathbf{y}^T\mathbf{y}}$ où $\mathbf{y} = \mathbf{Q}^T\mathbf{x} = (y_1, y_2, \dots, y_n)^T$

1. Formes quadratiques

Application : optimisation sous contrainte – quotient de Rayleigh

Mais : $$\mathbf{y}^T\mathbf{D}\mathbf{y} = (y_1, y_2, \dots, y_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} \lambda_i y_i^2$$

Et : $$\mathbf{y}^T\mathbf{y} = \sum_{i=1}^{n} y_i^2 \stackrel{\text{def}}{=} s$$

Donc :

$$R(\mathbf{x}) = \frac{\sum_{i=1}^{n} \lambda_i y_i^2}{s} = \sum_{i=1}^{n} p_i \lambda_i \qquad \text{avec} \quad p_i = \frac{y_i^2}{s} > 0$$

1. Formes quadratiques

Application : optimisation sous contrainte – quotient de Rayleigh

Le quotient de Rayleigh est donc le barycentre à poids positifs p_i des valeurs propres réelles λ_i de la matrice $\mathbf{A}$ associée à la forme quadratique $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$, c'est-à-dire :

$$R(\mathbf{x}) \in [\lambda_1, \lambda_n]$$

Le minimum de $R(\mathbf{x})$ est donc égal à la plus petite valeur propre (λ_1) et il est atteint en le vecteur propre $\mathbf{x}_1$ associé à λ_1 ; en effet :

$$R(\mathbf{x}_1) = \frac{{\mathbf{x}_1}^T\mathbf{A}\mathbf{x}_1}{{\mathbf{x}_1}^T\mathbf{x}_1} = \frac{{\mathbf{x}_1}^T\lambda_1\mathbf{x}_1}{{\mathbf{x}_1}^T\mathbf{x}_1} = \lambda_1$$

De même, le maximum de $R(\mathbf{x})$ est donc la plus grande valeur propre (λ_n) et il est atteint en le vecteur propre $\mathbf{x}_n$ associé à λ_n.

1. Formes quadratiques

Application : optimisation sous contrainte – quotient de Rayleigh

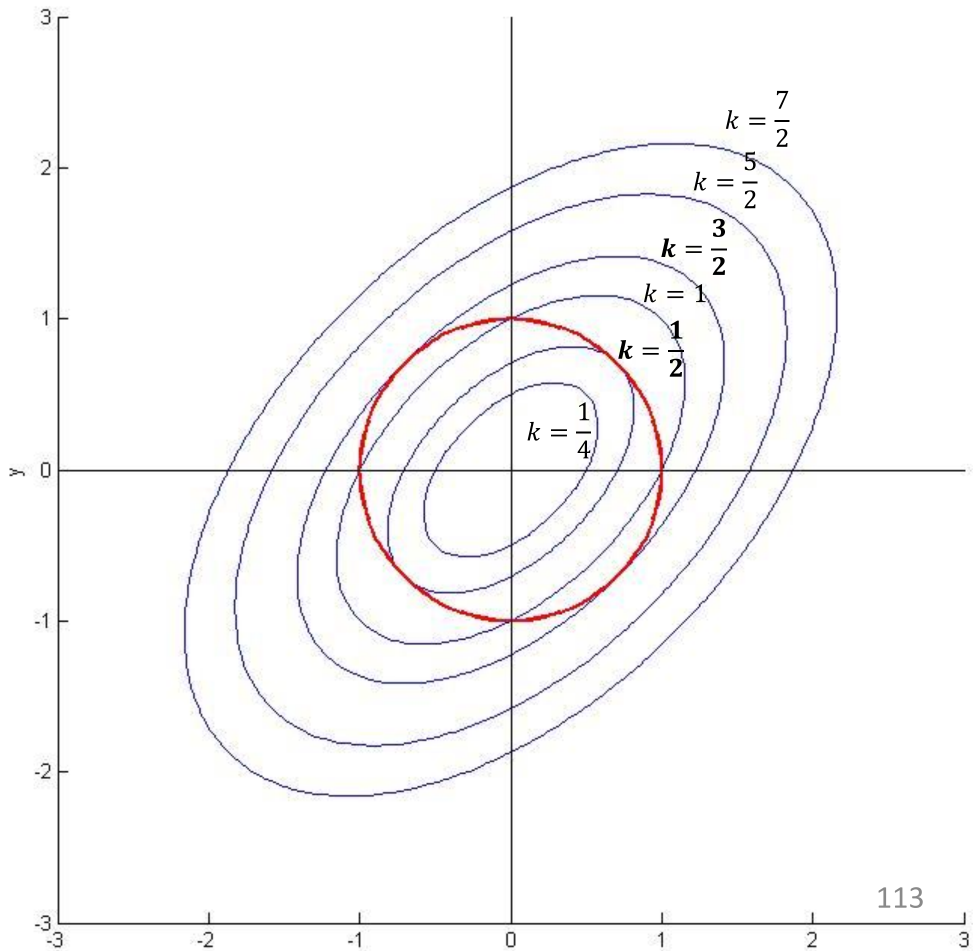
Exemple

Optimiser $f(x, y) = x^2 + y^2 - xy = k \in \mathbb{R}$ sous la contrainte $g(x, y) = x^2 + y^2 - 1 = 0$.

Les équations $x^2 + y^2 - xy = k$ forment une famille de courbes (plus précisément ici des ellipses).

Optimiser $x^2 + y^2 - xy = k$ sous la contrainte $g(x, y) = 0$ signifie chercher la plus petite (et/ou la plus grande) valeur de k telle qu'on ait simultanément :

$$\begin{cases} x^2 + y^2 - xy = k \\ x^2 + y^2 = 1 \end{cases}$$



1. Formes quadratiques

Application : optimisation sous contrainte – quotient de Rayleigh

On applique la méthode précédente utilisant le quotient de Rayleigh :

Théorème spectral :

1. Formes quadratiques

Application : optimisation sous contrainte – quotient de Rayleigh

- Le minimum de $f(x,y)$ est la valeur propre et il est atteint en le vecteur propre

  vérification : $f(\mathbf{x}_1) =$

- Le maximum de $f(x,y)$ est et il est atteint en

  vérification : $f(\mathbf{x}_2) =$

2. Polynômes vectoriels du 2nd ordre

Définitions

Exemple :

$$f(x,y,z) = x^2 + 2y^2 - z^2 + 8xy - 6xz + 12yz + 14x + 2y + 18z + 3$$

Donc :

2. Polynômes vectoriels du 2^{nd} ordre

Définitions

De façon générale un **polynôme vectoriel du 2^{nd} ordre** est défini par :

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c$$

où $\mathbf{x} = (x_1 \quad x_2 \quad \dots \quad x_n)^T \in \mathbb{R}^n$

$\mathbf{A} \in \mathbb{R}^{n\times n}$ symétrique

$\mathbf{b} = (b_1 \quad b_2 \quad \dots \quad b_n)^T \in \mathbb{R}^n$

$c \in \mathbb{R}$

2. Polynômes vectoriels du 2nd ordre

Définitions

Forme canonique

$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c$

Finalement :

$$f(\mathbf{x}) = \tilde{f}(\mathbf{x}_1) = \mathbf{x}_1^T\mathbf{A}\mathbf{x}_1 + s$$

avec $\mathbf{x_1} = \mathbf{x} + \mathbf{A}^{-1}\mathbf{b}$ et $s = c - \mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}$

Forme canonique d'un polynôme vectoriel du 2nd ordre

2. Polynômes vectoriels du 2nd ordre

Définitions

Exemple précédent :

$$f(\mathbf{x}) = x^2 + 2y^2 - z^2 + 8xy - 6xz + 12yz + 14x + 2y + 18z + 3$$

La forme canonique s'écrit alors :

$$f(\mathbf{x}) = \tilde{f}(\mathbf{x}_1) = \mathbf{x}_1^T \mathbf{A} \mathbf{x}_1 + s \qquad \text{avec} \qquad \mathbf{A} = \begin{pmatrix} 1 & 4 & -3 \\ 4 & 2 & 6 \\ -3 & 6 & -1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 7 \\ 1 \\ 9 \end{pmatrix}$$

$$\mathbf{x_1} = \mathbf{x} + \mathbf{A}^{-1}\mathbf{b}$$

$$s = c - \mathbf{b}^T \mathbf{A}^{-1} \mathbf{b}$$

On inverse $\mathbf{A}$, on trouve :

$$\mathbf{A}^{-1} = \frac{1}{92} \begin{pmatrix} 19 & 7 & -15 \\ 7 & 5 & 9 \\ -15 & 9 & 7 \end{pmatrix}$$

2. Polynômes vectoriels du 2^{nd} ordre

Définitions

Donc :

$$\mathbf{x_1} = \mathbf{x} + \mathbf{A}^{-1}\mathbf{b} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{1}{92}\begin{pmatrix} 19 & 7 & -15 \\ 7 & 5 & 9 \\ -15 & 9 & 7 \end{pmatrix}\begin{pmatrix} 7 \\ 1 \\ 9 \end{pmatrix} = \begin{pmatrix} x_1 = x + \dfrac{5}{92} \\ y_1 = y + \dfrac{135}{92} \\ z_1 = z + \dfrac{5}{92} \end{pmatrix}$$

$$\text{et} \quad s = c - \mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} = 3 - \begin{pmatrix} 7 & 1 & 9 \end{pmatrix}\frac{1}{92}\begin{pmatrix} 19 & 7 & -15 \\ 7 & 5 & 9 \\ -15 & 9 & 7 \end{pmatrix}\begin{pmatrix} 7 \\ 1 \\ 9 \end{pmatrix} = \cdots = \frac{61}{92}$$

Finalement :

$$\tilde{f}(\mathbf{x}_1) = \begin{pmatrix} x_1 & y_1 & z_1 \end{pmatrix}\begin{pmatrix} 1 & 4 & -3 \\ 4 & 2 & 6 \\ -3 & 6 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \frac{61}{92}$$

$$= x_1^2 + 2y_1^2 - z_1^2 + 8x_1y_1 - 6x_1z_1 + 12y_1z_1 + \frac{61}{92}$$

Les termes d'ordre 1 (les termes en x, y et z) ont disparu.

2. Polynômes vectoriels du 2nd ordre

Définitions

Remarque

Dans le cas d'un polynôme du 2nd degré à une variable on a : $f(x) = ax^2 + 2bx + c$

$$f(x) = a\left(x^2 + \frac{2b}{a}x + \frac{c}{a}\right)$$

$$= a\left[\left(x + \frac{b}{a}\right)^2 - \frac{b^2}{a^2} + \frac{c}{a}\right]$$

$$= a\left[\left(\left(x + \frac{b}{a}\right)^2 - \frac{b^2 - ac}{a^2}\right)\right]$$

$$= ax_1^2 + s \qquad \text{avec} \quad x_1 = x + \frac{b}{a} \quad \text{et} \quad s = -\frac{b^2 - ac}{a} = c - \frac{b^2}{a}$$

Comme x_1, a et b sont des réels sont peut écrire :

$$f(x) = \tilde{f}(x_1) = x_1^T a x_1 + s \quad \text{avec} \quad x_1 = x + ba^{-1} \quad \text{et} \quad s = c - b^T a^{-1} b$$

On voit alors l'analogie avec le cas général.

2. Polynômes vectoriels du 2nd ordre

Gradient

Le gradient d'un polynôme vectoriel $f(\mathbf{x})$ avec $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ est défini par :

$$\overrightarrow{\text{grad}} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right)^T$$

Or :

$$f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c = \underbrace{\sum_{k=1}^{n}\sum_{j=1}^{n} a_{kj}x_k x_j}_{\mathbf{x}^T\mathbf{A}\mathbf{x}} + 2\underbrace{\sum_{k=1}^{n} b_k x_k}_{\mathbf{b}^T\mathbf{x}} + c$$

$\mathbf{A} = (a_{ij})$ *symétrique*

2. Polynômes vectoriels du 2nd ordre

Gradient

Donc :

$$\frac{\partial f}{\partial x_i} = 2a_{ii}x_i + \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j + \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ji}x_j + 2b_i = 2a_{ii}x_i + 2\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j + 2b_i = 2\sum_{j=1}^{n} a_{ij}x_j + 2b_i$$

Finalement :

$$\overrightarrow{\text{grad}} f = 2\begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{pmatrix} + 2\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = 2\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + 2\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$= 2\mathbf{Ax} + 2\mathbf{b}$$

2. Polynômes vectoriels du 2nd ordre

Gradient

A retenir

$$\overrightarrow{\text{grad}}(\mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c) = 2\mathbf{A}\mathbf{x} + 2\mathbf{b}$$

Cette formule est la généralisation de : $\dfrac{d}{dx}(ax^2 + 2bx + c) = 2ax + 2b$

Exemple précédent

$$f(\mathbf{x}) = x^2 + 2y^2 - z^2 + 8xy - 6xz + 12yz + 14x + 2y + 18z + 3$$

$$= \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c$$

$$\text{avec} \quad \mathbf{A} = \begin{pmatrix} 1 & 4 & -3 \\ 4 & 2 & 6 \\ -3 & 6 & -1 \end{pmatrix}, \qquad \mathbf{b} = (7 \quad 1 \quad 9)^T \quad \text{et} \quad c = 3$$

2. Polynômes vectoriels du 2nd ordre

Gradient

Donc :

$$\overrightarrow{\mathrm{grad}}\big(f(\mathbf{x})\big) = 2\mathbf{A}\mathbf{x} + 2\mathbf{b} = 2\begin{pmatrix} x + 4y - 3z \\ 4x + 2y + 6z \\ -3x + 6y - z \end{pmatrix} + 2\begin{pmatrix} 7 \\ 1 \\ 9 \end{pmatrix} = \begin{pmatrix} 2x + 8y - 6z + 14 \\ 8x + 4y + 12z + 2 \\ -6x + 12y - 2z + 18 \end{pmatrix}$$

2. Polynômes vectoriels du 2nd ordre

Optimisation

On cherche à optimiser (maximiser ou minimiser) un polynôme vectoriel du 2nd ordre. Cela revient à déterminer le ou les vecteurs $\mathbf{x}$ tels que :

$$\overrightarrow{\text{grad}}\big(f(\mathbf{x})\big) = 2\mathbf{A}\mathbf{x} + 2\mathbf{b} = \mathbf{0}$$

L'optimum est donc atteint pour :

$$\mathbf{x}_{\text{opt}} = -\mathbf{A}^{-1}\mathbf{b}$$

Et la valeur de l'optimum est :

$$\begin{aligned} \boldsymbol{f}\big(\mathbf{x}_{\text{opt}}\big) &= \mathbf{x}_{\text{opt}}{}^{T}\mathbf{A}\mathbf{x}_{\text{opt}} + 2\mathbf{b}^{T}\mathbf{x}_{\text{opt}} + c \\ &= \big(-\mathbf{A}^{-1}\mathbf{b}\big)^{T}\mathbf{A}\big(-\mathbf{A}^{-1}\mathbf{b}\big) + 2\mathbf{b}^{T}\big(-\mathbf{A}^{-1}\mathbf{b}\big) + c \\ &= \mathbf{b}^{T}\mathbf{A}^{-1}\mathbf{A}\mathbf{A}^{-1}\mathbf{b} - 2\mathbf{b}^{T}\mathbf{A}^{-1}\mathbf{b} + c \end{aligned}$$

Finalement :

$$\boldsymbol{f}\big(\mathbf{x}_{\text{opt}}\big) = c - \mathbf{b}^{T}\mathbf{A}^{-1}\mathbf{b} \quad \text{complément de Schur de } \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^{T} & c \end{pmatrix}$$

2. Polynômes vectoriels du 2nd ordre

Optimisation

A retenir

Si $\mathbf{A}$ *est une matrice symétrique alors le polynôme vectoriel du 2nd degré* $\mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c$ *admet un optimum en* $\mathbf{x}_{opt} = -\mathbf{A}^{-1}\mathbf{b}$ *et cet optimum est égal à* $s = c - \mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}$, *complément de Schur de la matrice* $\begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & c \end{pmatrix}$.

Ces formules sont la généralisation de :

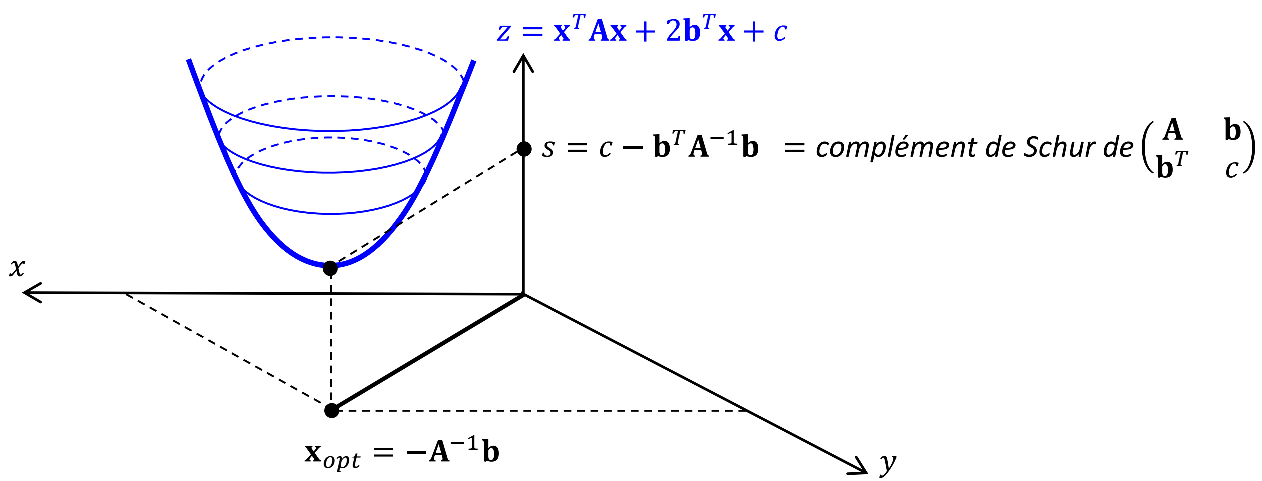
$$\frac{d}{dx}\underbrace{(ax^2 + 2bx + c)}_{f(x)} = 2ax + 2b = 0 \quad \Rightarrow \quad x_{opt} = -\frac{b}{a} \quad \text{et} \quad f(x_{opt}) = a\left(-\frac{b}{a}\right)^2 + 2b\left(-\frac{b}{a}\right) + c = c - \frac{b^2}{a}$$

2. Polynômes vectoriels du 2nd ordre

Optimisation

Remarque

En 2D, c'est-à-dire pour des polynômes vectoriels du 2nd degré à 2 variables x et y, on peut illustrer les résultats précédents de la façon suivantes :



2. Polynômes vectoriels du 2nd ordre

Optimisation

Exemple précédent

$$f(\mathbf{x}) = x^2 + 2y^2 - z^2 + 8xy - 6xz + 12yz + 14x + 2y + 18z + 3$$

$$= \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{b}^T\mathbf{x} + c$$

$$\text{avec} \quad \mathbf{A} = \begin{pmatrix} 1 & 4 & -3 \\ 4 & 2 & 6 \\ -3 & 6 & -1 \end{pmatrix}, \qquad \mathbf{b} = (7 \quad 1 \quad 9)^T \quad \text{et} \quad c = 3$$

Donc :

$$\mathbf{x}_{opt} = -\mathbf{A}^{-1}\mathbf{b} = -\frac{1}{92}\begin{pmatrix} 19 & 7 & -15 \\ 7 & 5 & 9 \\ -15 & 9 & 7 \end{pmatrix}\begin{pmatrix} 7 \\ 1 \\ 9 \end{pmatrix} = -\frac{1}{92}\begin{pmatrix} 5 \\ 135 \\ 5 \end{pmatrix}$$

Et :

$$f(\mathbf{x}_{opt}) = s = c - \mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} = \frac{61}{92}$$