

# Optimisation

## Chapitre 2 : Formes quadratiques

Joseph GERGAUD, Serge GRATTON & Daniel RUIZ

10 septembre 2024

Le but de ce chapitre 2 d'étudier les **formes quadratiques dans  $\mathbb{R}^n$**

### Motivations

- La fonction  $f$  dans le problème aux moindres carrées linéaires est une forme quadratique généralisée.

$$(P) \begin{cases} \text{Min } f(\beta) = \frac{1}{2} \|y - X\beta\|^2 \\ \beta \in \mathbb{R}^p. \end{cases}$$

- Le développement limité à l'ordre 2 d'une fonction à valeurs dans  $\mathbb{R}$  est une forme quadratique généralisée.

**Définition 2.1.1 – Formes bilinéaires**

Soit  $E$  un espace vectoriel réel de dimension finie. On appelle forme bilinéaire sur  $E$ , toute application  $f$  de  $E \times E$  dans  $\mathbb{R}$  vérifiant les propriétés suivantes, pour tous vecteurs  $\mathbf{u}$ ,  $\tilde{\mathbf{u}}$ ,  $\mathbf{v}$ , et  $\tilde{\mathbf{v}}$  de  $E$  et tout scalaire  $\lambda$  de  $\mathbb{R}$  :

$$\begin{aligned}f(\mathbf{u} + \tilde{\mathbf{u}}, \mathbf{v}) &= f(\mathbf{u}, \mathbf{v}) + f(\tilde{\mathbf{u}}, \mathbf{v}) & f(\lambda \mathbf{u}, \mathbf{v}) &= \lambda f(\mathbf{u}, \mathbf{v}) \\f(\mathbf{u}, \mathbf{v} + \tilde{\mathbf{v}}) &= f(\mathbf{u}, \mathbf{v}) + f(\mathbf{u}, \tilde{\mathbf{v}}) & f(\mathbf{u}, \lambda \mathbf{v}) &= \lambda f(\mathbf{u}, \mathbf{v})\end{aligned}$$

$f$  est en fait linéaire par rapport à chacune de ses deux variables.

**Définition 2.1.2 – Formes bilinéaires symétrique****Forme bilinéaire symétrique**

Soit  $E$  un espace vectoriel réel de dimension finie, et soit  $f$  une forme bilinéaire sur  $E$ . On dit que  $f$  est symétrique si, pour tous vecteurs  $\mathbf{x}$  et  $\mathbf{y}$  de  $E$ , on a :

$$f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x}).$$

Soit  $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  une base de  $E$ . Toute forme bilinéaire  $f$  est entièrement déterminée par la connaissance des réels  $f(\mathbf{e}_i, \mathbf{e}_j)$ , pour tout  $1 \leq i, j \leq n$ . En effet, soient

$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$  et  $\mathbf{y} = \sum_{i=1}^n y_i \mathbf{e}_i$  deux vecteurs de  $E$ . Par linéarité à gauche, et à droite, on peut écrire, après développement :

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

Introduisons alors  $\mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\mathbf{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ , les vecteurs de  $\mathbb{R}^n$  formés des composantes de  $\mathbf{x}$  et  $\mathbf{y}$  dans la base  $\mathcal{B}$ , et  $\mathbf{A}$  la matrice des coefficients  $f(\mathbf{e}_i, \mathbf{e}_j)$ ,

$$\mathbf{A} = \begin{pmatrix} f(\mathbf{e}_1, \mathbf{e}_1) & \dots & f(\mathbf{e}_1, \mathbf{e}_n) \\ \vdots & \ddots & \vdots \\ f(\mathbf{e}_n, \mathbf{e}_1) & \dots & f(\mathbf{e}_n, \mathbf{e}_n) \end{pmatrix}.$$

En utilisant ces notations, on peut alors écrire la valeur de  $f(\mathbf{x}, \mathbf{y})$  en terme du produit matriciel suivant :

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{X}^T \mathbf{A} \mathbf{Y}.$$

**Proposition 2.1.3**

Si  $f$  est une forme bilinéaire symétrique sur  $E$ , alors la matrice associée à  $f$  dans une base quelconque de  $E$  est symétrique.

**Exemple 2.1.1.** Exemple dans  $\mathbb{R}^3$

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= x_1y_1 + 2x_2y_2 + 3x_3y_3 + x_1y_3 + x_3y_1 + x_2y_3 \\ &= (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}. \end{aligned}$$

**Définition 2.2.1 – Formes quadratiques**

On appelle forme quadratique associée à la forme bilinéaire  $f$ , l'application  $q$  définie de  $E$  dans  $\mathbb{R}$  par :

$$\forall \mathbf{x} \in E, \quad q(\mathbf{x}) = f(\mathbf{x}, \mathbf{x}).$$

**Remarque 2.2.1.**

- On a aussi, en utilisant la matrice  $\mathbf{A}$  de  $f$  dans une base  $\mathcal{B}$  de  $E$  :

$$q(\mathbf{x}) = \mathbf{X}^T \mathbf{A} \mathbf{X},$$

où  $\mathbf{X}$  est le vecteur des coordonnées de  $\mathbf{x}$  dans la base  $\mathcal{B}$ . Ainsi,  $\mathbf{A}$  représente aussi la matrice de la forme quadratique  $q$  dans la base  $\mathcal{B}$ .

- Par contre, la représentation matricielle d'une forme quadratique n'est pas unique. En effet, pour une forme quadratique donnée, il existe plusieurs formes bilinéaires qui peuvent lui être associées.

**Exemple 2.2.1.** Exemple dans  $\mathbb{R}^3$

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= x_1 y_1 - 2x_2 y_2 + 3x_3 y_3 + x_1 y_3 + x_3 y_1 + 4x_2 y_3 + 4x_3 y_2 \\ &= (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 4 \\ 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}. \end{aligned}$$

La forme quadratique associée est

$$q(\mathbf{x}) = x_1^2 - 2x_2^2 + 3x_3^2 + 2x_1 x_3 + 8x_2 x_3 \quad \text{soit} \quad q(\mathbf{x}) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 4 \\ 1 & 4 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Mais on a aussi, du point de vue matriciel :

$$q(\mathbf{x}) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 2 \\ 0 & -2 & 8 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 2 & 8 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

**Remarque 2.2.2.**

- Pour un vecteur  $\mathbf{u} \in E$  donné,  $q(\mathbf{u})$  est un polynôme homogène de degré 2. Ainsi, tout polynôme homogène de degré 2 par rapport aux coordonnées d'un vecteur  $\mathbf{u}$  de  $E$  peut correspondre à une forme quadratique  $q$ .
- En outre, à la question "*existe-t-il une forme bilinéaire symétrique dont  $q$  soit la forme quadratique et si oui, est-elle unique ?*", la réponse est "*oui*".

Voici comment procéder : il suffit pour cela d'écrire la matrice  $\mathbf{A} = (a_{ij})$  associée à ce polynôme homogène de degré 2 en plaçant, sur la diagonale, les coefficients  $a_{ii}$  correspondant aux termes en  $x_i^2$ , et sur les termes hors diagonaux  $a_{ij}$  et  $a_{ji}$  la moitié des coefficients des termes en  $x_i x_j$ .

- Enfin, si à une même forme quadratique  $q$ , on peut effectivement associer diverses formes bilinéaires  $f$  (de matrice associée  $\mathbf{A}_f$  dans une base  $\mathcal{B}$  fixée), ces formes bilinéaires ont toutes en commun **la même partie symétrique** :

$$s(\mathbf{u}, \mathbf{v}) = \frac{f(\mathbf{u}, \mathbf{v}) + f(\mathbf{v}, \mathbf{u})}{2}, \quad \text{de matrice associée } \frac{\mathbf{A}_f + \mathbf{A}_f^T}{2} \quad \text{indépendante de } f.$$



**Exemple 2.2.2.** Exemple dans  $\mathbb{R}^3$  :

$$q(\mathbf{x}) = 5x_1^2 + 12x_2^2 - 6x_3^2 - 8x_2x_3 + 5x_3x_1 - x_2x_1,$$

la forme matricielle symétrique associée étant

$$q(\mathbf{x}) = (x_1, x_2, x_3) \begin{pmatrix} 5 & -1/2 & 5/2 \\ -1/2 & 12 & -4 \\ 5/2 & -4 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Soit  $f$  une forme bilinéaire symétrique sur  $E$ , et  $q$  la forme quadratique associée. Pour tous vecteurs  $\mathbf{u}$  et  $\mathbf{v}$  de  $E$  et tout scalaire  $\lambda$ , on a :

- $q(\lambda \mathbf{u}) = f(\lambda \mathbf{u}, \lambda \mathbf{u}) = \lambda^2 q(\mathbf{u})$  :  $q$  n'est pas linéaire.
- $f(\mathbf{u}, \mathbf{v}) = \frac{1}{4} (q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u} - \mathbf{v}))$ .
- $f(\mathbf{u}, \mathbf{v}) = \frac{1}{2} (q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v}))$ .
- Pour une forme quadratique  $q$  donnée, la forme bilinéaire symétrique  $f$  qui lui est associée est aussi appelée forme polaire de  $q$ .
- $q$  est dite semi-définie positive ssi  $\forall \mathbf{x} \in E, q(\mathbf{x}) \geq 0$ .
- $q$  est dite semi-définie négative ssi  $-q$  est semi-définie positive.
- $q$  est dite indéfinie ssi  $q$  n'est ni semi-définie positive, ni semi-définie négative.
- $q$  est dite définie positive si  $\forall \mathbf{x} \in E, q(\mathbf{x}) \geq 0$  et  $q(\mathbf{x}) = 0 \Rightarrow \mathbf{x} = \mathbf{0}$ .

**Définition 2.3.1 – Produit scalaire**

On rappelle que un **produit scalaire** sur un  $\mathbb{R}$ -espace vectoriel  $E$  est une forme **bilinéaire, symétrique, et définie positive**. La définie positivité d'une forme bilinéaire  $f$  sur  $E$  correspond en fait à la définie positivité de sa forme quadratique, à savoir :

$$\forall \mathbf{u} \in E, q(\mathbf{u}) = f(\mathbf{u}, \mathbf{u}) \geq 0 \quad \text{et} \quad q(\mathbf{u}) = f(\mathbf{u}, \mathbf{u}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}.$$

**Proposition 2.3.2**

Soit  $E$  un  $\mathbb{R}$ -espace vectoriel de dimension finie, et soit  $q$  une forme quadratique définie positive sur  $E$ . Alors, la forme polaire de  $q$ , qui est une forme bilinéaire symétrique (ou à symétrie hermitienne si le corps de référence est  $\mathbb{C}$ ) définie positive sur  $E$ , constitue un produit scalaire sur  $E$ , et pour la norme associée,  $E$  est un espace EUCLIDIEN. On notera  $\langle \mathbf{u}, \mathbf{v} \rangle$  le produit scalaire.

**Exemple 2.3.1.** Dans  $\mathbb{R}^3$ , soit la forme quadratique  $q$  définie par

$$q(\mathbf{u}) = x^2 + 6xy + 4yz + 14y^2 + z^2,$$

avec  $\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ . Voyons si  $q$  est définie positive. Pour ce faire, décomposons  $q$  en somme de trois carrés dans  $\mathbb{R}$  :

$$q(\mathbf{u}) = (x + 3y)^2 + 5\left(y + \frac{2}{5}z\right)^2 + \frac{1}{5}z^2.$$

Cette somme de carrés dans  $\mathbb{R}$  est positive, donc la forme quadratique  $q$  est semi-définie positive ( $\forall \mathbf{u} \in E, q(\mathbf{u}) \geq 0$ ). De plus :

$$\begin{aligned} q(\mathbf{u}) = (x + 3y)^2 + 5\left(y + \frac{2}{5}z\right)^2 + \frac{1}{5}z^2 = 0 &\Leftrightarrow \begin{cases} x + 3y = 0 \\ y + \frac{2}{5}z = 0 \\ z = 0 \end{cases} \\ &\Leftrightarrow x = y = z = 0 \\ &\Leftrightarrow \mathbf{u} = \mathbf{0}. \end{aligned}$$

**Bilan :** cette forme quadratique est bien définie positive, et la forme bilinéaire symétrique associée

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_3 + 2x_3y_2 + 14x_2y_2 + x_3y_3 \\ &= (x_1, x_2, x_3) \begin{pmatrix} 1 & 3 & 0 \\ 3 & 14 & 2 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \end{aligned}$$

définit bien un produit scalaire sur  $\mathbb{R}^3$ .

**Proposition 2.4.1**

On démontre les résultats suivants :

- Tout endomorphisme symétrique d'un espace euclidien est diagonalisable.
- Ses valeurs propres sont réelles.
- Les espaces propres sont deux à deux orthogonaux.
- Il existe toujours une base orthonormée formée de vecteurs propres.

**Remarque 2.4.1.**

- Le fait que, dans un espace euclidien, tout endomorphisme symétrique se diagonalise dans une base orthonormale de vecteurs propres s'écrit en termes d'algèbre linéaire sous la forme :

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T, \text{ avec } \mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I} \text{ et } \mathbf{\Lambda} = \text{diag}(\lambda_i)_{1 \leq i \leq n}.$$

C'est d'ailleurs l'un des principaux intérêts des notations matricielles, à savoir d'exprimer de manière très concise des propriétés ou des transformations.

- Soit  $q$  une forme quadratique sur un  $\mathbb{R}$ -espace vectoriel euclidien  $E$  et  $f$  sa forme bilinéaire symétrique associée. Soit  $\mathbf{A}$  la matrice symétrique des coefficients  $f(\mathbf{e}_i, \mathbf{e}_j)$ , où les  $\mathbf{e}_k$  sont les vecteurs de la base canonique par exemple.

$$\forall \mathbf{x}, \mathbf{y} \in E, \quad f(\mathbf{x}, \mathbf{y}) = \mathbf{X}^T \mathbf{A} \mathbf{Y},$$

$\mathbf{X}$  et  $\mathbf{Y}$  étant les vecteurs des composantes de  $\mathbf{x}$  et  $\mathbf{y}$  dans la base  $\mathcal{B} = (\mathbf{e}_k)_{1 \leq k \leq n}$ .

- La matrice  $\mathbf{A}$  étant symétrique, elle est diagonalisable dans une base orthonormée de vecteurs propres  $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$ , avec  $\mathbf{U}^T = \mathbf{U}^{-1}$
- Dans la base de vecteurs propres la forme quadratique s'écrit alors

$$\forall \mathbf{x} \in E, \quad q(\mathbf{x}) = \mathbf{Z}^T \mathbf{\Lambda} \mathbf{Z} = \sum_{i=1}^n \lambda_i z_i^2,$$

où les  $z_i$ ,  $i = 1, \dots, n$ , sont les composantes de  $\mathbf{x}$  dans la base des vecteurs propres :

$$\mathbf{x} = \sum_{i=1}^n z_i \mathbf{u}_i.$$

Cette dernière égalité peut aussi s'écrire matriciellement sous la forme :

$$\mathbf{X} = \mathbf{U} \mathbf{Z} \Leftrightarrow \mathbf{Z} = \mathbf{U}^T \mathbf{X}.$$

- Il est à noter que  $z_i = \mathbf{u}_i^T \mathbf{X}$  n'est rien d'autre que le produit scalaire du  $i^{\text{ème}}$  vecteur propre de  $\mathbf{A}$  (i.e. la  $i^{\text{ème}}$  colonne de  $\mathbf{U}$ ) avec le vecteur  $\mathbf{x}$ . Cela correspond au calcul des composantes d'un vecteur dans une base orthonormée donnée, que l'on obtient effectivement par produit scalaire avec les vecteurs de cette base.
- D'un point de vue géométrique, l'écriture de  $q$  sous la forme

$$\sum_{i=1}^n \lambda_i z_i^2$$

signifie simplement que la forme quadratique  $q$  se décompose en paraboles élémentaires, dirigées selon les axes des vecteurs propres  $\mathbf{u}_i$ , et de courbures respectives  $\lambda_i$ .



- De manière équivalente, on peut aussi dire que les iso-contours

$$q(\mathbf{x}) = C^{\text{ste}}$$

sont des coniques dans  $\mathbb{R}^n$  dont les axes principaux correspondent aux vecteurs propres de la matrice  $\mathbf{A}$  associée à la forme quadratique  $q$ .

- **Cas particulier :** si la forme quadratique  $q$  est définie positive, alors les valeurs propres  $\lambda_i$  ci-dessus sont nécessairement toutes strictement positives, et les iso-contours  $q(\mathbf{x}) = C^{\text{ste}}$  correspondent alors à des hyper-ellipsoïdes dans  $\mathbb{R}^n$ . Par exemple,  $\lambda_1 z_1^2 + \lambda_2 z_2^2 = C$ , avec  $\lambda_1 > 0$  et  $\lambda_2 > 0$ , est l'équation d'une ellipse dans  $\mathbb{R}^2$ , et l'équation

$$\lambda_1 z_1^2 + \lambda_2 z_2^2 + \lambda_3 z_3^2 = C,$$

avec  $\lambda_{1,2,3}$  strictement positifs, représenterait une surface dans  $\mathbb{R}^3$  du type "*ballon de rugby*".

La figure ci dessous illustre la forme géométrique d'une nappe quadratique, à savoir le dessin dans  $\mathbb{R}^3$  d'une forme quadratique de  $\mathbb{R}^2$  dans  $\mathbb{R}$ , où  $(x, y)$  jouent le rôle de  $\mathbf{x} \in \mathbb{R}^2$  et  $z = \frac{1}{2}\mathbf{x}^T \mathbf{A} \mathbf{x}$  (avec  $\mathbf{A}$  matrice  $2 \times 2$  symétrique définie positive).

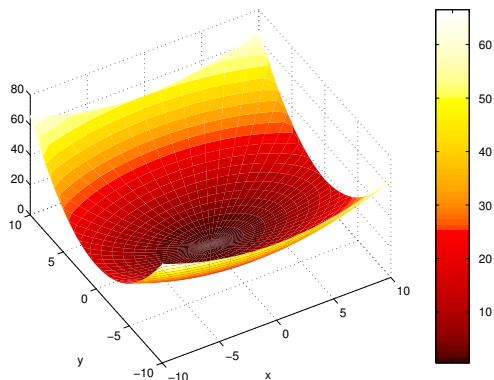


FIGURE 1 – Un exemple de forme quadratique en dimension 2

**Définition 2.5.1 – Fonctionnelle quadratique généralisée**

On appelle **Fonctionnelle quadratique généralisée** toute application  $f$  de  $\mathbb{R}^n$  dans  $\mathbb{R}$  sous la forme :

$$\forall \mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{v}^T \mathbf{x} + c,$$

où  $\mathbf{A}$  est une matrice de  $\mathcal{M}_n(\mathbb{R})$ ,  $\mathbf{v}$  un vecteur de  $\mathbb{R}^n$ , et  $c$  une constante réelle.

On appelle **terme quadratique** associé à la fonctionnelle  $f$  le terme  $\frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x}$ .

**Remarque 2.5.1.** On peut toujours se ramener au cas où la matrice  $\mathbf{A}$  est symétrique, car on a :

$$\forall \mathbf{u} \in \mathbb{R}^n, \mathbf{u}^T \mathbf{A} \mathbf{u} = \mathbf{u}^T \left( \frac{\mathbf{A} + \mathbf{A}^T}{2} \right) \mathbf{u}.$$

$$Q = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad A = Q \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} Q^T, \quad \mathbf{v} = (1 \ 2)^T.$$

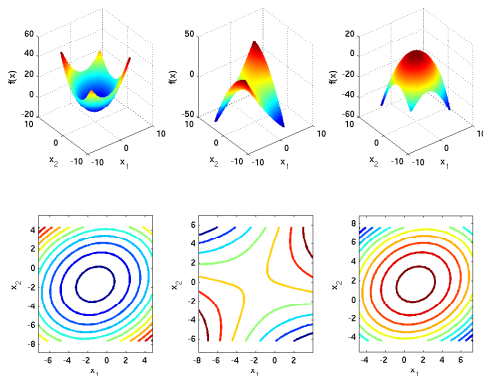


FIGURE 2 – Cas où le rang de  $\mathbf{A}$  est 2. De gauche à droite  $\lambda_1 = 1, \lambda_2 = 3/2$  (la matrice  $\mathbf{A}$  est définie positive);  $\lambda_1 = 1, \lambda_2 = -3/2$  (la matrice  $\mathbf{A}$  est indéfinie);  $\lambda_1 = -1, \lambda_2 = -3/2$  (la matrice  $\mathbf{A}$  est définie négative).  $\theta = \pi/6$ .

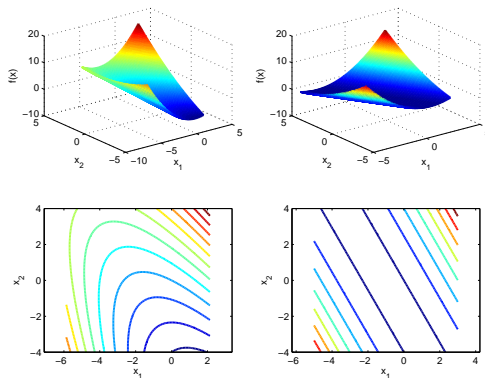


FIGURE 3 – Cas où le rang de  $\mathbf{A}$  est 1 ;  $\lambda_1 = 1, \lambda_2 = 0$  et  $\mathbf{v} = \begin{pmatrix} 1 & 2 \end{pmatrix}^T$  dans le premier cas, et  $\mathbf{v} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \end{pmatrix}^T$  dans le deuxième cas (le vecteur  $\mathbf{b}$  est dans l'image de la matrice  $\mathbf{A}$ ),  $\theta = \pi/6$ .