Distance et approximation

Moindres carrés

MAT-2930 Algèbre linéaire appliquée

Jean-François Lalonde

Image

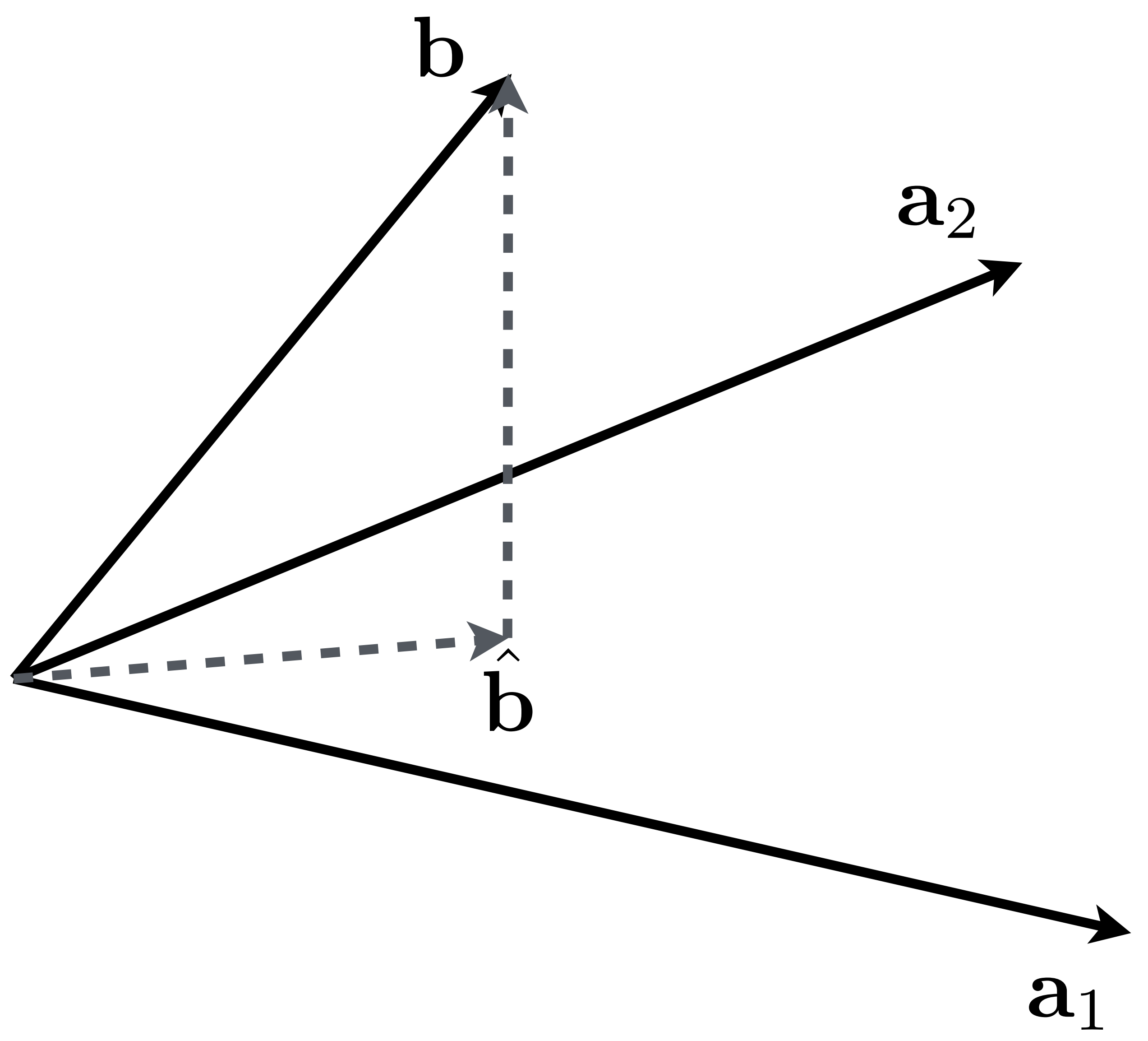
Moindres carrés

On nomme la solution à l'équation

$$\mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$$

On nomme cette équation l'**équation normale**

la solution aux moindres carrés
car c'est elle qui minimise l'erreur quadratique moyenne



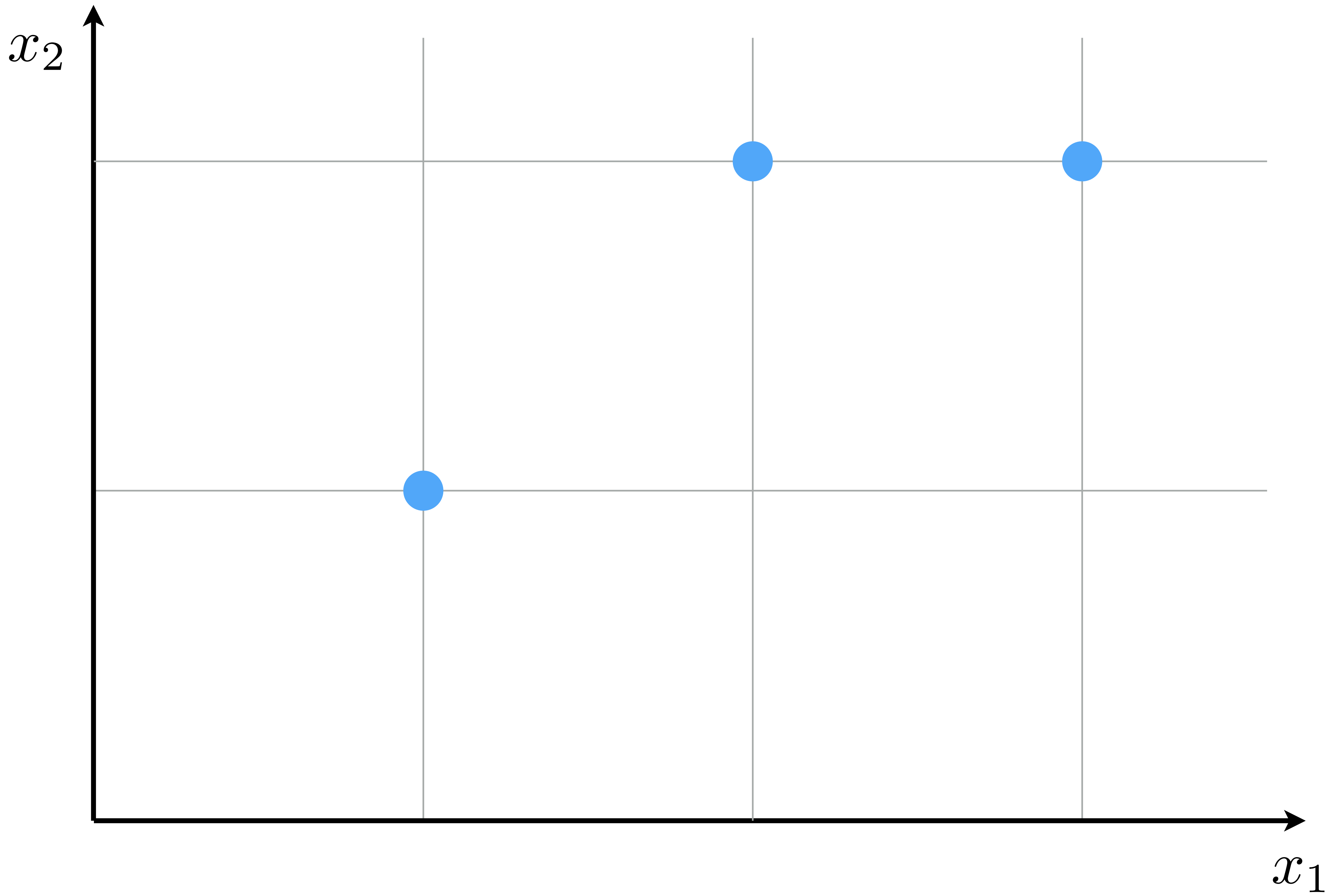
$$\epsilon = ||\mathbf{b} - \hat{\mathbf{b}}||$$

Exemple : droite

Trouvons la meilleure droite qui approxime les trois points suivants.

Écrivons les équations

$$x_2 = C + Dx_1$$



Exprimons le problème sous forme matricielle

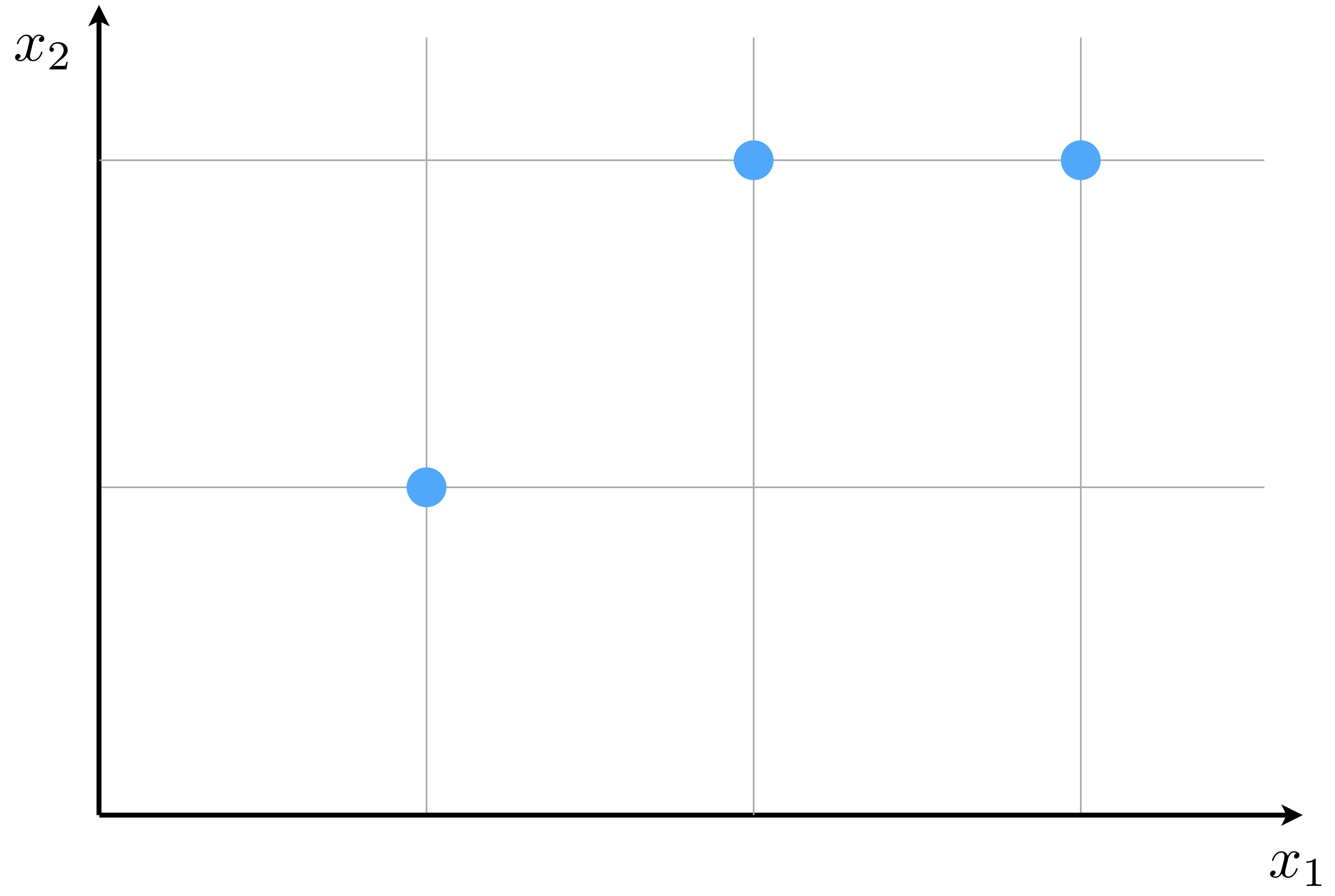
Exemple : droite

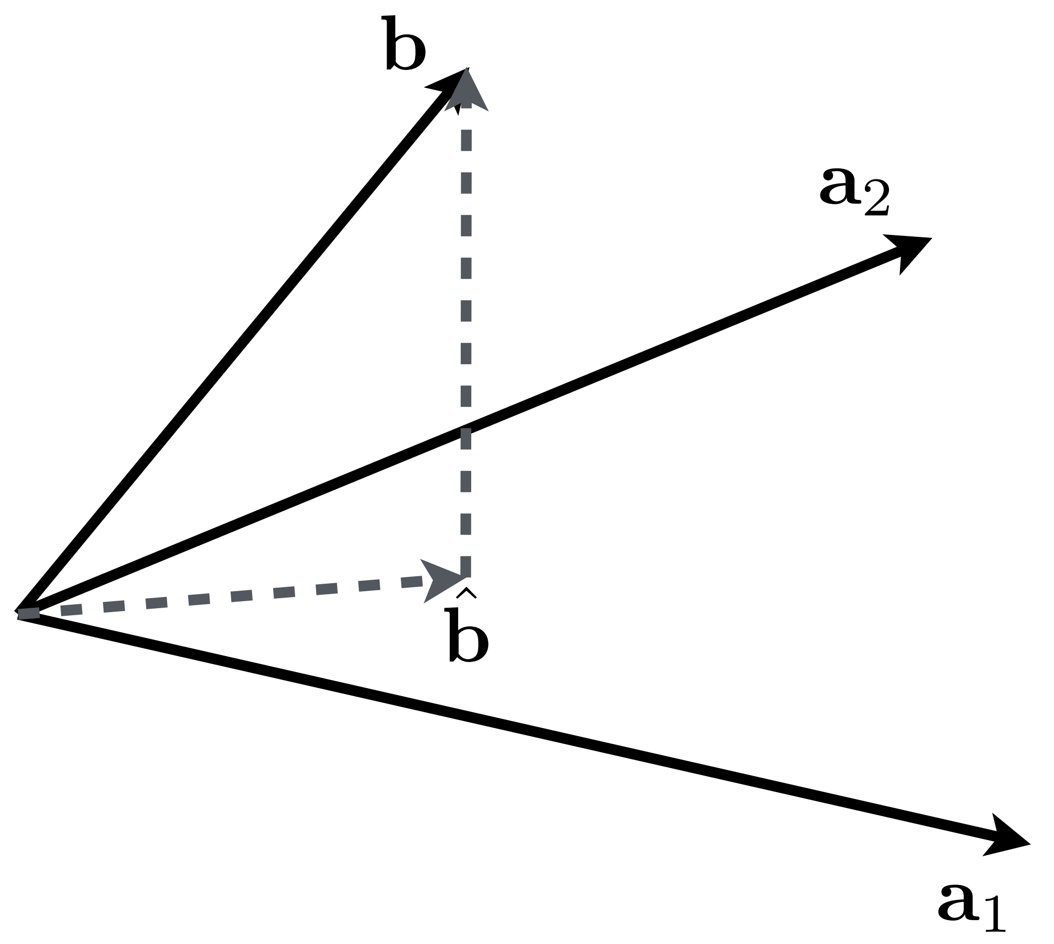
Trouvons la meilleure droite qui approxime les trois points suivants.

À quoi cette équation correspond-elle ?

$$x_2 = C + Dx_1$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$





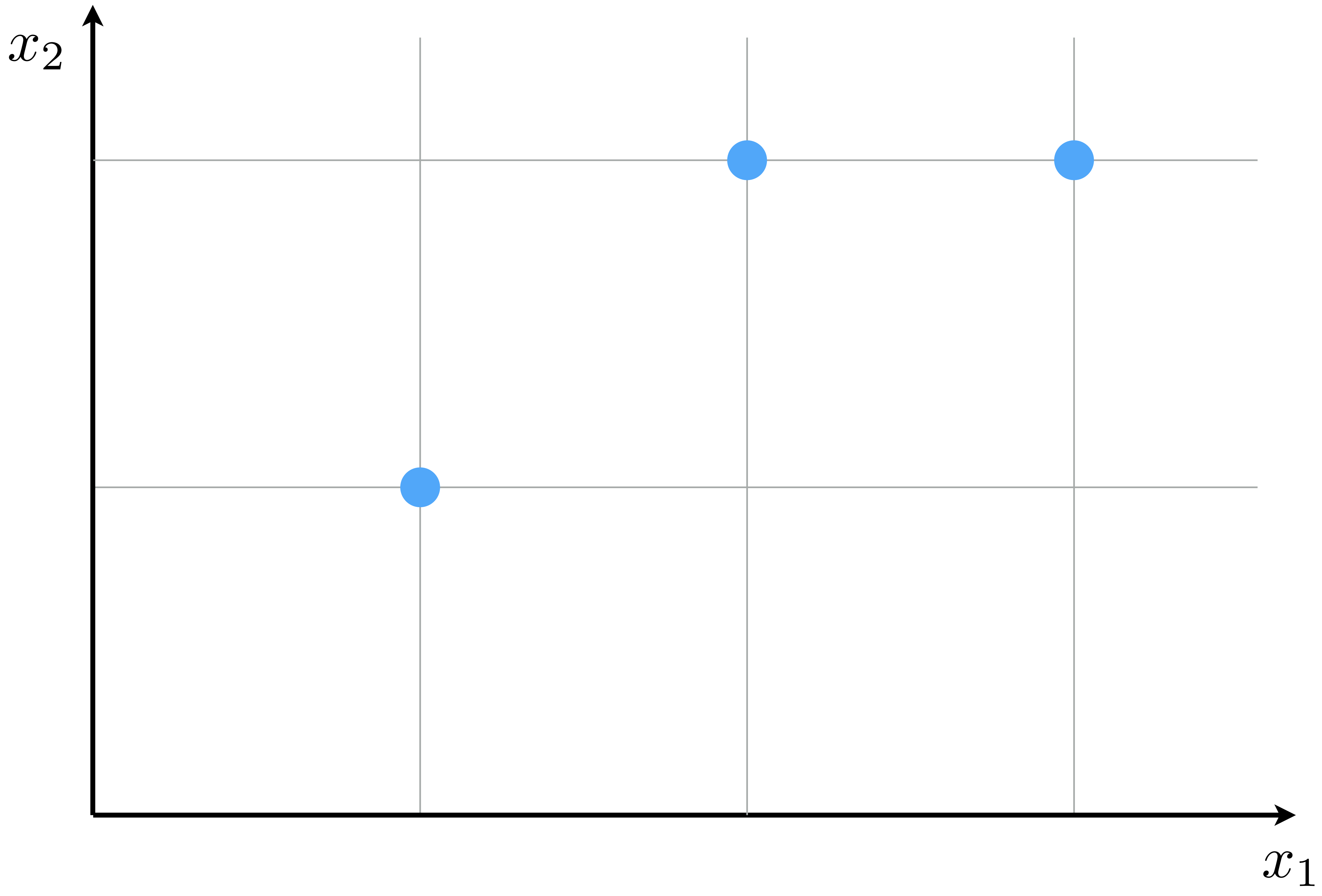
Exemple : droite

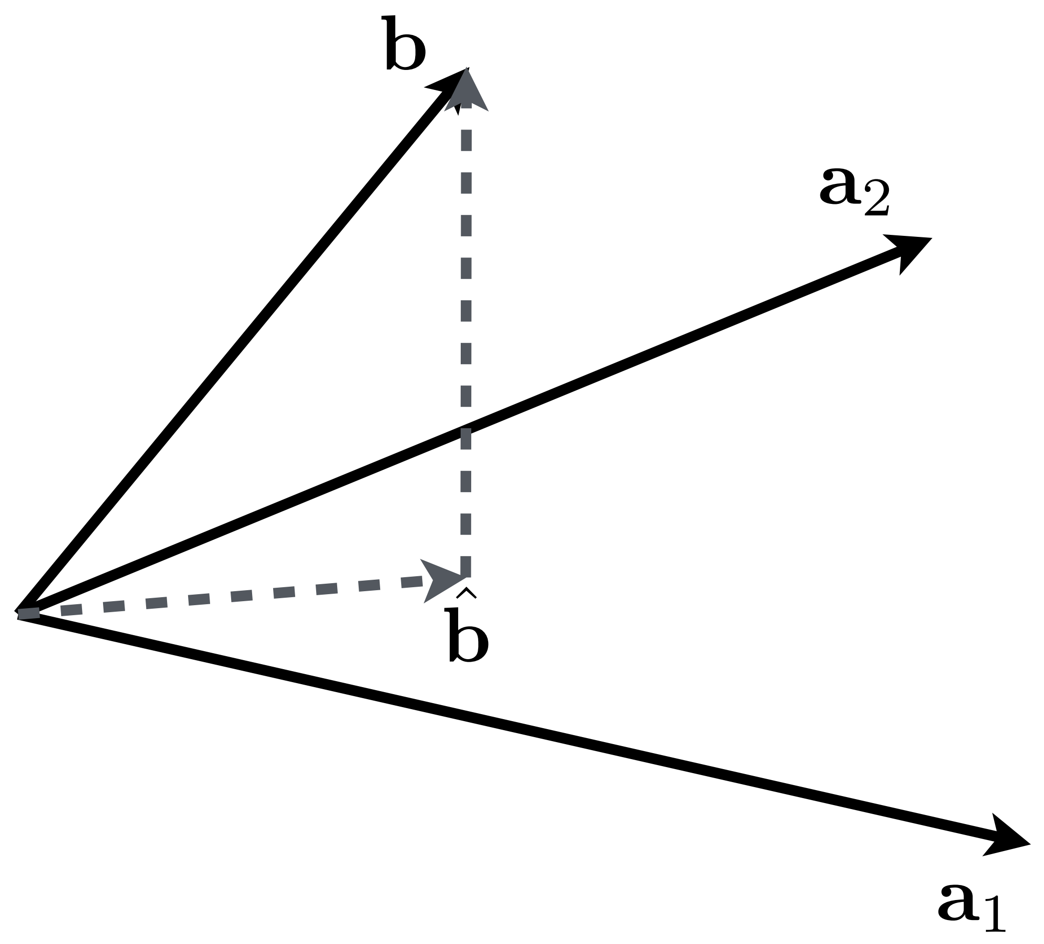
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Trouvons la meilleure droite qui approxime les trois points suivants.

Quelle erreur minimise-t-elle ?

$$x_2 = C + Dx_1$$





Dans quels sous-espaces sont situés ces vecteurs ?

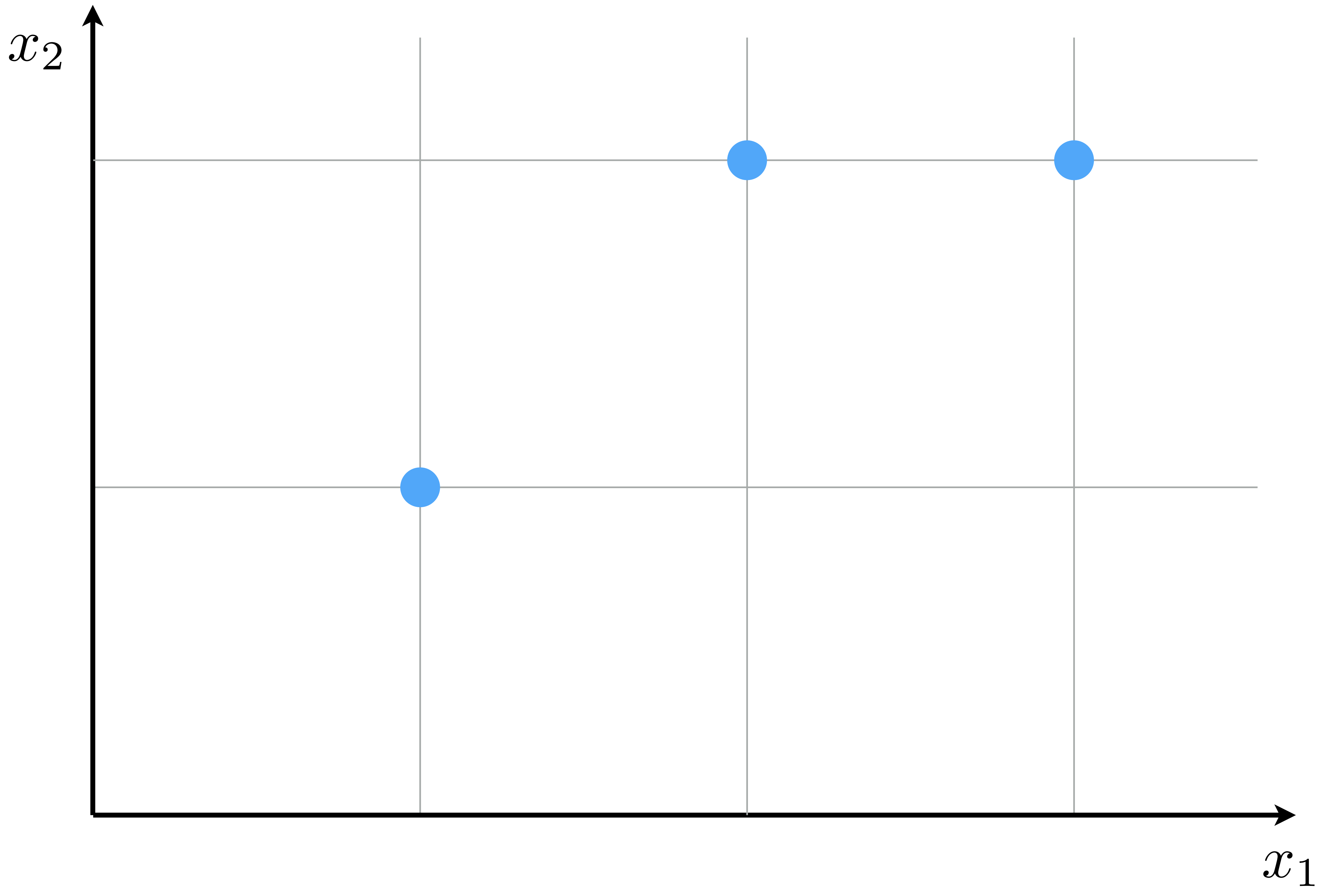
Exemple : droite

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Trouvons la meilleure droite qui approxime les trois points suivants.

Quelles valeurs de C et D minimisent cette erreur ?

$$x_2 = C + Dx_1$$



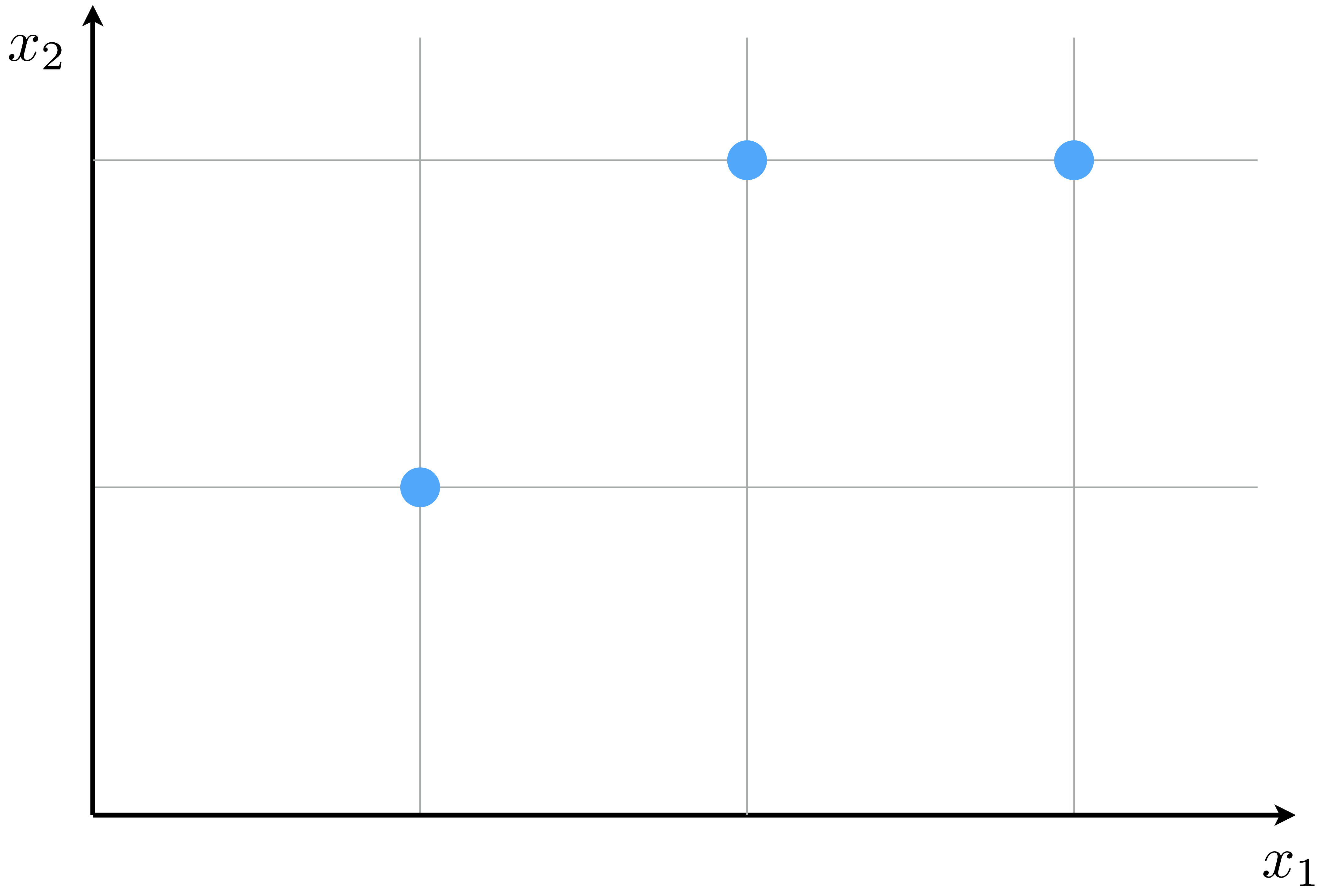
Exemple : droite

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Trouvons la meilleure droite qui approxime les trois points suivants.

Quelle est l'erreur de cette approximation ?

$$x_2 = \frac{2}{3} + \frac{1}{2} x_1$$



Inversibilité

$$\mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$$

$$\hat{\mathbf{x}} = (\mathbf{A}^\top \mathbf{A})^{-1}\mathbf{A}^\top \mathbf{b}$$

C'est la matrice pseudo-inverse

Si les colonnes de $\mathbf{A}$ sont linéairement indépendantes, alors $\mathbf{A}^\top\mathbf{A}$ est inversible.

Preuve ? Commençons avec $\mathbf{A}^\top\mathbf{A}\mathbf{x} = \mathbf{0}$.

Factorisation QR $\mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$

Peut-on utiliser la factorisation QR pour résoudre l'équation normale ci-haut ?

Orthogonalité

$$\mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$$

Si les colonnes de $\mathbf{A}$ sont orthonormales :

Orthogonalité

$$\mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$$

Si les colonnes de $\mathbf{A}$ sont orthogonales :

Orthogonalité

$$\mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$$

Si les colonnes de $\mathbf{A}$ sont orthogonales : $\hat{\mathbf{x}} = \mathbf{D}^{-1}\mathbf{A}^\top \mathbf{b}$

$$\mathbf{A} = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Orthogonalité

$$\mathbf{A}^\top \mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top \mathbf{b}$$

Si les colonnes de $\mathbf{A}$ sont orthogonales, la projection du vecteur $\mathbf{b}$ dans l'espace des colonnes de $\mathbf{A}$ nous donne directement la solution.

$$\mathbf{A} = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$