

6.3 Familles orthogonales et projections orthogonales

Définition 65 (famille orthogonale).

Une famille de vecteurs $\{\vec{u}_1, \dots, \vec{u}_p\}$ de $\mathbb{R}^n$ est dite *orthogonale* si

Exemples

Théorème 60. *Soit $\{\vec{u}_1, \dots, \vec{u}_p\}$ une famille orthogonale de $\mathbb{R}^n$ formée de vecteurs $\vec{u}_i$ non nuls pour tout $1 \leq i \leq p$. Alors la famille*

Preuve

Remarque

Définition 66 (base orthogonale).

Soit W un sous-espace vectoriel de $\mathbb{R}^n$. On appelle *base orthogonale* de W toute base qui est composée d'une famille orthogonale.

Exemples

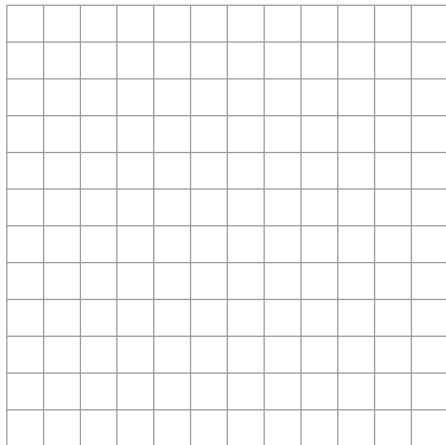
Application

Pour trouver les composantes d'un vecteur dans une base orthogonale, on peut utiliser soit les techniques de changement de base ou de résolution de système (échelonnement d'une matrice augmentée) classiques, soit utiliser les propriétés du produit scalaire :

Théorème 61. *Soient W un sous-espace vectoriel de $\mathbb{R}^n$ et $(\vec{u}_1, \dots, \vec{u}_p)$ une base orthogonale de W . Alors*

Preuve

Projections orthogonales



Exemple

Définition 67 (famille orthonormale).

On dira qu'une famille $\{\vec{u}_1, \dots, \vec{u}_p\}$ de vecteurs de $\mathbb{R}^n$ est *orthonormale* si

1. la famille est orthogonale,
2. tous les vecteurs de la famille sont des vecteurs unitaires.

On dira alors que $(\vec{u}_1, \dots, \vec{u}_p)$ est une *base orthonormale* de

Exemples

Théorème 62. Soit $U = (\vec{u}_1 \cdots \vec{u}_n) \in M_{m \times n}(\mathbb{R})$ une matrice. Alors les colonnes de U sont orthonormales si et seulement si

Preuve

Exemple

Théorème 63. Soit $U \in M_{m \times n}(\mathbb{R})$ une matrice dont les colonnes sont orthonormales. Soient $\vec{v}, \vec{w} \in \mathbb{R}^n$. Alors

- 1.
- 2.
- 3.

Remarques

1. Les propriétés ci-dessus montrent que l'application linéaire $\vec{v} \mapsto U\vec{v}$ conserve les longueurs et l'orthogonalité.
2. On verra le lien entre le produit scalaire de deux vecteurs et l'angle formé par ces derniers dans une série d'exercices.

Généralisation de la projection à $\mathbb{R}^n$

Soit W un sous-espace vectoriel de $\mathbb{R}^n$ ainsi qu'un vecteur $\vec{v} \notin W$. Alors la projection de $\vec{v}$ sur W , notée $\text{proj}_W(\vec{v})$ est

1. l'unique vecteur de W tel que
2. l'unique vecteur de W qui

Théorème 64. *Soit W un sous-espace vectoriel de $\mathbb{R}^n$. Alors tout vecteur $\vec{v} \in \mathbb{R}^n$ peut s'écrire de façon unique comme la somme de deux vecteurs*

Plus précisément, si $W = \text{span}\{\vec{u}_1, \dots, \vec{u}_p\}$ où $(\vec{u}_1, \dots, \vec{u}_p)$ est une base orthogonale de W , on a

Interprétation géométrique

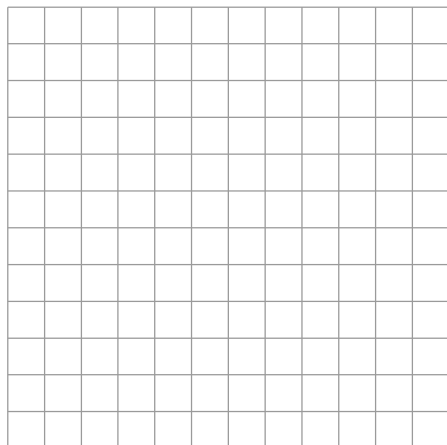
Remarques

1. Le vecteur $\text{proj}_W(\vec{v})$ ne dépend pas du choix de la base orthogonale de W .
2. Si la base $(\vec{u}_1, \dots, \vec{u}_p)$ est orthonormale, alors

Exemple

Théorème 65 (de la meilleure approximation). *Soient W un sous-espace vectoriel de $\mathbb{R}^n$ et $\vec{v} \in \mathbb{R}^n$. Soit $\text{proj}_W(\vec{v})$ la projection orthogonale de $\vec{v}$ sur W . Alors $\text{proj}_W(\vec{v})$ est la meilleure approximation de $\vec{v}$ dans W , autrement dit, on a*

Illustration



Théorème 66. *Soit W un sous-espace vectoriel de $\mathbb{R}^n$ et supposons que $(\vec{u}_1, \dots, \vec{u}_p)$ est une base orthonormale de W . Alors pour tout $\vec{v} \in \mathbb{R}^n$, on a*

Preuve

Remarques

Définition 68 (matrice orthogonale).

Une matrice $U \in M_{n \times n}(\mathbb{R})$ inversible telle que

Remarques

6.4 Procédé de Gram-Schmidt

Le procédé de Gram-Schmidt permet d'obtenir une base orthogonale (ou même orthonormée) pour tout sous-espace vectoriel W de $\mathbb{R}^n$. Une telle base est nécessaire pour pouvoir calculer la projection orthogonale $\text{proj}_W(\vec{v})$ d'un vecteur $\vec{v} \in \mathbb{R}^n$ sur W .

Exemple d'orthogonalisation d'une base

Théorème 67 (de Gram-Schmidt). *Soient W un sous-espace vectoriel de $\mathbb{R}^n$ et $(\vec{v}_1, \dots, \vec{v}_p)$ une base de W . Alors on peut construire une base orthogonale de W selon le procédé suivant :*

Exemple

6.5 Factorisation QR

Si les colonnes d'une matrice $A \in M_{m \times n}(\mathbb{R})$ sont linéairement indépendantes, on peut leur appliquer le procédé de Gram-Schmidt. Cela revient à trouver une factorisation de la matrice A :

Théorème 68 (Factorisation QR). *Soit $A \in M_{m \times n}(\mathbb{R})$ une matrice dont les colonnes sont linéairement indépendantes. Alors il existe une factorisation $A = QR$ où*

1.

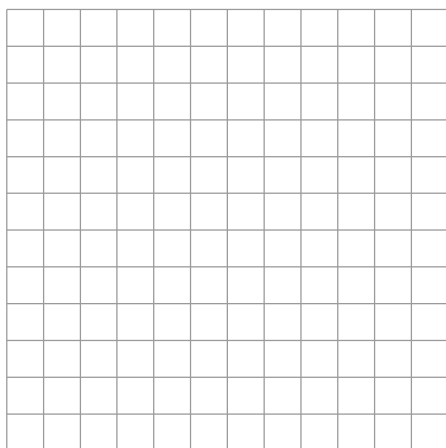
2.

Remarques

Exemple

6.6 Méthode des moindres carrés

Exemple



Définition 69 (solution au sens des moindres carrés).

Soient $A \in M_{m \times n}(\mathbb{R})$ une matrice et $\vec{b} \in \mathbb{R}^m$. On appelle *solution au sens des moindres carrés* de $A\vec{x} = \vec{b}$ le vecteur

Démarche

Suite de l'exemple

L'équation normale

Théorème 69. *L'ensemble des solutions au sens des moindres carrés de $A\vec{x} = \vec{b}$ est égal à l'ensemble des solutions de*

On appelle cette équation l'équation normale.

Remarque

Exemples

Théorème 70. *Soit $A \in M_{m \times n}(\mathbb{R})$ une matrice. Alors les propriétés suivantes sont équivalentes*

1.

2.

3.

Lien avec la factorisation QR

Théorème 71. *Soit $A \in M_{m \times n}(\mathbb{R})$ une matrice dont les colonnes sont linéairement indépendantes. Considérons la décomposition $A = QR$. Alors pour tout $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ admet une unique solution au sens des moindres carrés donnée par*

6.7 Droite de régression

Supposons qu'on dispose d'un certain nombre de données expérimentales du type $(x_1, y_1), \dots, (x_n, y_n)$ et qu'on cherche à établir une formule permettant de prédire les valeurs d'une variable en fonction de l'autre variable. Ce type de situation est étudié en analyse statistique. On résout souvent le problème à l'aide d'une méthode qui fait appel aux moindres carrés.

On s'intéresse ici à la relation la plus élémentaire entre deux variables x et y , relation du premier degré du type $y = ax + b$. Lorsqu'on dispose d'un nuage de points $(x_1, y_1), \dots, (x_n, y_n)$, on cherche à déterminer les paramètres a et b qui rendent la droite $y = ax + b$ aussi proche que possible des points expérimentaux.

Exemple

Remarques