10. Bases orthonormales et procédé de Gram-Schmidt

Section 4.4

MTH1007

J. Guérin, N. Lahrichi, S. Le Digabel
Polytechnique Montréal

H2023

(v2)

Plan

Vecteur orthonormaux

- Les vecteurs $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ de $\mathbb{R}^m$ (avec $n \leq m$) sont *orthonormaux* si

$$\mathbf{q}_i^\top \mathbf{q}_j = \begin{cases} 0 & \text{lorsque } i \neq j \quad \text{(orthogonalité)} \\ 1 & \text{lorsque } i = j \quad \text{(vecteurs \textit{unitaires})} \end{cases}$$

  pour tous les $i, j \in \{1, 2, \ldots, n\}$

- Géométriquement, les vecteurs sont de longueur 1 et perpendiculaires entre eux

Matrices orthogonales (1/2)

- Si les colonnes de la matrice $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \in \mathbb{R}^{m \times n}$ sont orthonormales alors $Q^\top Q = I \in \mathbb{R}^{n \times n}$

- La matrice Q n'est pas nécessairement carrée : $Q^\top Q = I \in \mathbb{R}^{n \times n}$ mais $QQ^\top \in \mathbb{R}^{m \times m} \neq I$

- Dans le cas où Q est carrée alors on dit que c'est une *matrice orthogonale* et on a $Q^\top Q = QQ^\top = I$, et donc $Q^{-1} = Q^\top$

- Les colonnes d'une matrice orthogonale forment une base de $\mathbb{R}^n$. C'est une *base orthonormale*

Matrices orthogonales (2/2)

- Une matrice orthogonale $Q \in \mathbb{R}^{n \times n}$ préserve le produit scalaire : Pour tous vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^n$, on a

$$(Q\mathbf{x})^\top (Q\mathbf{y}) = \mathbf{x}^\top Q^\top Q \mathbf{y} = \mathbf{x}^\top \mathbf{y}$$

- Géométriquement, ceci signifie que les longueurs et les angles sont préservés

- $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ pour tout $\mathbf{x} \in \mathbb{R}^n$

Matrices orthogonales : Exemples

- La matrice orthogonale la plus évidente est l'**identité** : $Q = I$
- Matrices de **rotation** : $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in \mathbb{R}^{2\times 2}$
- Les matrices de **permutation** sont orthogonales
- Matrices de **réflexion** : Si $\mathbf{u} \in \mathbb{R}^n$ unitaire, on définit Q par la *transformation de Householder* :

$$Q = I - 2\mathbf{u}\mathbf{u}^\top \in \mathbb{R}^{n\times n}$$

Q est symétrique, orthogonale, et $Q^2 = I$. Noter aussi que $\mathbf{u}$ ne reste pas comme une colonne de Q

Exemples avec $\mathbf{u}_1 = (1, 0)$ et $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, -1)$

Projection : Cas $n \leq m$

- Si les vecteurs $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ de $\mathbb{R}^m$ sont orthonormaux alors la matrice de projection sur le sous-espace de $\mathbb{R}^m$ de dimension n engendré par ces vecteurs ($C(Q)$) se simplifie et devient
$$P = Q(Q^\top Q)^{-1}Q^\top = QQ^\top \in \mathbb{R}^{m\times m}$$
- $\hat{\mathbf{x}} = (Q^\top Q)^{-1}Q^\top \mathbf{b} = Q^\top \mathbf{b} \in \mathbb{R}^n$ avec $\mathbf{b} \in \mathbb{R}^m$: La solution au sens des moindres carrés devient très facile à calculer : Plus besoin de $(A^\top A)^{-1}$
- $\mathbf{p} = Q\hat{\mathbf{x}} = QQ^\top \mathbf{b} = P\mathbf{b} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^\top \mathbf{b} \\ \mathbf{q}_2^\top \mathbf{b} \\ \cdots \\ \mathbf{q}_n^\top \mathbf{b} \end{bmatrix} =$
$$\mathbf{q}_1(\mathbf{q}_1^\top \mathbf{b}) + \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{b}) + \ldots + \mathbf{q}_n(\mathbf{q}_n^\top \mathbf{b}) \in \mathbb{R}^m$$

Projection : Cas $n = m$

- $C(Q) = \mathbb{R}^n$ et la projection d'un vecteur $\mathbf{b} \in \mathbb{R}^n$ sur $C(Q)$ est lui-même : $P = I$

- De plus, on a $\mathbf{p} = \mathbf{b}$ et

$$\mathbf{b} = \mathbf{q}_1 \left(\mathbf{q}_1^\top \mathbf{b}\right) + \mathbf{q}_2 \left(\mathbf{q}_2^\top \mathbf{b}\right) + \cdots + \mathbf{q}_n \left(\mathbf{q}_n^\top \mathbf{b}\right) \in \mathbb{R}^n$$

Ceci est la décomposition de $\mathbf{b}$ dans la base des $\mathbf{q}_j$, $j \in \{1, 2, \ldots, n\}$

- **Exemple :**

$$\begin{aligned} \begin{bmatrix} 2 \\ 3 \end{bmatrix} &= 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \tfrac{5}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \tfrac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \end{aligned}$$

Procédé de Gram-Schmidt (1/2)

Soit $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ des vecteurs linéairement indépendants de $\mathbb{R}^m$ et W le sous-espace engendré par ces vecteurs ($n \leq m$).
La procédure suivante produit une base orthonormale de W :

(1) $\mathbf{v}_1 = \mathbf{u}_1$

(2) $\mathbf{v}_2 = \mathbf{u}_2 - \dfrac{\mathbf{v}_1^\top \mathbf{u}_2}{\mathbf{v}_1^\top \mathbf{v}_1} \mathbf{v}_1$ ($\mathbf{v}_2 \perp \mathbf{v}_1$ car obtenu par $\mathbf{u}_2 - \mathbf{p}$ avec $\mathbf{p}$ la projection de $\mathbf{u}_2$ sur $\mathbf{v}_1$)

(3) $\mathbf{v}_3 = \mathbf{u}_3 - \dfrac{\mathbf{v}_1^\top \mathbf{u}_3}{\mathbf{v}_1^\top \mathbf{v}_1} \mathbf{v}_1 - \dfrac{\mathbf{v}_2^\top \mathbf{u}_3}{\mathbf{v}_2^\top \mathbf{v}_2} \mathbf{v}_2$

⋮

Procédé de Gram-Schmidt (2/2)

(j) $$\mathbf{v}_j = \mathbf{u}_j - \sum_{k=1}^{j-1} \frac{\mathbf{v}_k^\top \mathbf{u}_j}{\mathbf{v}_k^\top \mathbf{v}_k} \mathbf{v}_k$$

⋮

(n) $$\mathbf{v}_n = \mathbf{u}_n - \sum_{k=1}^{n-1} \frac{\mathbf{v}_k^\top \mathbf{u}_n}{\mathbf{v}_k^\top \mathbf{v}_k} \mathbf{v}_k$$

(n + 1) $\mathbf{q}_j = \dfrac{\mathbf{v}_j}{||\mathbf{v}_j||}$ pour tout $j \in \{1, 2, \ldots, n\}$

Les vecteurs $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ forment la matrice Q et la base orthonormale recherchée. On a donc $W = C(Q) = C(A)$ si $A = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_n]$

Procédé de Gram-Schmidt : Remarques

- Les “nouveaux” $\mathbf{q}_i$ sont orthogonaux aux “vieux” $\mathbf{u}_j$ (avec $i > j$) :
  - $\mathbf{q}_2 \perp \mathbf{u}_1$
  - $\mathbf{q}_3 \perp \mathbf{u}_1$
  - $\mathbf{q}_3 \perp \mathbf{u}_2$
  - $\ldots$
- Si $m = n$, on obtient une base orthonormale de $\mathbb{R}^n$
- Le coût du procédé est en $\mathcal{O}(mn^2)$
- La transformation de Householder permet aussi d’obtenir une base orthornormale, à moindre coût
- La forme matricielle du procédé de Gram-Schmidt est la décomposition QR

Factorisation QR (1/2)

- Si $A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ et $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ est obtenue par le procédé de Gram-Schmidt alors les $\mathbf{u}_i$ sont des combinaisons linéaires des $\mathbf{q}_i$ (et vice versa), de sorte que

$$A = QR \text{ avec } R \in \mathbb{R}^{n\times n}$$

- R est triangulaire supérieure et peut s'obtenir par $R = Q^\top A$
- Ses éléments diagonaux correspondent aux normes des $\mathbf{v}_j$ du procédé de Gram-Schmidt

Factorisation QR (2/2)

- La décomposition fonctionne pour des matrices carrées ou rectangulaires ($n \leq m$), à condition que les colonnes de A soient indépendantes

- Illustration avec $m = n = 3$:

$$\begin{bmatrix} \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ \\ \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^\top \mathbf{u}_1 & \mathbf{q}_1^\top \mathbf{u}_2 & \mathbf{q}_1^\top \mathbf{u}_3 \\ & \mathbf{q}_2^\top \mathbf{u}_2 & \mathbf{q}_2^\top \mathbf{u}_3 \\ & & \mathbf{q}_3^\top \mathbf{u}_3 \end{bmatrix}$$

- Les formules des moindres carrés se simplifient en $\hat{\mathbf{x}} = R^{-1} Q^\top \mathbf{b}$ où R^{-1} se calcule facilement par remontée triangulaire