Norme et fonction norme de Frobenius

Franck Jeannot

Montréal, Canada, Avril 2023, AA805, v1.0

Abstract

A summary of the Frobenius norm that explains its origins, applications, and different equation variations.

Keywords: Norm, Frobenius norm, trace, transposition, produit scalaire, apprentissage automatique, cybersécurité, Jacobienne, cryptographie

1. Introduction

Une *fonction norme* est une fonction mathématique qui prend un vecteur ou une matrice en entrée et retourne un nombre non négatif qui mesure la taille ou la longueur de ce vecteur ou de cette matrice.

La *fonction norme de Frobenius* est une fonction mathématique qui permet de mesurer par exemple la **similarité** entre deux vecteurs en utilisant leur **produit scalaire** [1]. La norme de Frobenius d'une matrice est donnée par la racine carrée de la somme des carrés de tous les éléments de la matrice. Si on suppose la matrice suivante : $A = \begin{pmatrix} 1 & 2 & 3 4 & 5 & 6 7 & 8 & 9 \end{pmatrix}$, alors :

$$\|A\|_F = \sqrt{\sum i = 1^n \sum_{j=1}^{m} |a_{ij}|^2} \tag{1}$$

$$\|A\|_F = \sqrt{|1|^2 + |2|^2 + |3|^2 + |4|^2 + |5|^2 + |6|^2 + |7|^2 + |8|^2 + |9|^2} = \sqrt{285} \tag{2}$$

Ainsi, la norme [2] de Frobenius de cette matrice 3x3 est d'environ 16,88.

1. Un produit scalaire est une opération mathématique qui associe deux vecteurs pour donner un scalaire. Un scalaire est simplement un nombre, contrairement à un vecteur qui a une direction et une magnitude. Le produit scalaire mesure la projection d'un vecteur sur un autre et permet de calculer l'angle entre eux

2. Les barres verticales entourant la matrice A dans la formule $|A|_F = x$ indiquent la norme de Frobenius de la matrice A.

2. Usage (fonction norme de Frobenius)

La fonction norme de Frobenius est couramment utilisée en analyse numérique, traitement du signal et **apprentissage automatique**[1] [2] pour mesurer la similarité entre des vecteurs de données ou de caractéristiques. Elle est utile pour résoudre des problèmes de classification, de clustering ou de réduction de dimensionnalité, ainsi que dans l'algèbre linéaire, la théorie de l'information et le traitement du signal.

3. Origine (fonction norme de Frobenius)

La **fonction norme de Frobenius** est nommée d'après le mathématicien allemand Ferdinand Georg Frobenius [3], qui l'a introduite dans son article *"Über Matrizen aus positiven Elementen"* [3] en 1908 [4]. Frobenius a utilisé cette fonction pour définir une norme sur les matrices qui ont des coefficients positifs, et il a montré que cette norme est subordonnée à la norme de vecteur de colonne euclidienne. La fonction norme de Frobenius est également appelée **norme euclidienne de matrice**, **norme de Hilbert-Schmidt**, **norme L2 de matrice** ou **norme matricielle** [5].

4. La fonction norme de Frobenius

La norme de Frobenius d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est définie [4] par :

$$\|A\|_F := (\operatorname{tr} A^*A)^{1/2} = (\operatorname{tr} AA^*)^{1/2} = \sqrt{\sum_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} |A_{ij}|^2} \tag{4}$$

où A^* désigne la matrice adjointe de A et *tr* la **trace** [5]. On peut donc écrire :

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{i,j}|^2} \tag{5}$$

$$\|A\|_F^2 = \operatorname{tr}\left(AA^T\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{i,j}|^2. \tag{6}$$

---

3. Dans cet article, Frobenius étudie des matrices carrées composées d'éléments positifs, qui sont utilisées dans divers domaines des mathématiques appliquées. Il montre que ces matrices peuvent être factorisées de manière unique en un produit de matrices de la forme BC, où B est une matrice triangulaire supérieure avec des éléments positifs sur la diagonale, et C est une matrice triangulaire inférieure avec des éléments positifs sur la diagonale. Cette factorisation peut être utilisée pour déterminer les valeurs et les vecteurs propres de ces matrices, ainsi que pour résoudre des systèmes d'équations linéaires. Il présente également des applications de ces matrices dans la théorie des nombres, la géométrie analytique, et la théorie des groupes.

4. sur $\mathrm{M}_{m,n}(K)^3$ est celle qui dérive du produit scalaire ou hermitien standard sur cet espace, à savoir :

$$(A,B) \in \mathrm{M}_{m,n}(K)^2 \mapsto \langle A,B\rangle = \operatorname{tr}(A^*B) = \operatorname{tr}(BA^*) \tag{3}$$

5. https://fr.wikipedia.org/wiki/Trace_(alg%C3%A8bre)

C'est une expression de la norme de Frobenius d'une matrice A de dimensions $n \times n$. La norme de Frobenius est définie comme **la racine carrée de la somme des carrés de tous les éléments de la matrice**. L'équation montre que la norme de Frobenius peut également être exprimée en termes de la *trace* de la matrice AA^T, qui est la somme des carrés des éléments diagonaux de la matrice AA^T. La **trace** d'une matrice est **la somme des éléments diagonaux de la matrice**.

5. Norme de Frobenius et détails

Pour deux vecteurs colonnes $\boldsymbol{x}$ et $\boldsymbol{y}$ de dimension n, leur produit scalaire est donné par $\boldsymbol{x}^\top \boldsymbol{y}$, où $\top$ désigne l'opération de transposition. La norme de Frobenius de $\boldsymbol{x}$ est définie comme étant la racine carrée de la somme des carrés de ses éléments, c'est-à-dire $|\boldsymbol{x}|_2 = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$.

En utilisant ces deux définitions, la fonction norme de Frobenius mesure la similarité entre deux vecteurs $\boldsymbol{x}$ et $\boldsymbol{y}$ en prenant leur produit scalaire et en le divisant par le produit de leurs normes on a :

$$\boldsymbol{x}^\top \boldsymbol{y} = \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2 \cos\theta \tag{7}$$

où θ est l'angle entre les vecteurs $\boldsymbol{x}$ et $\boldsymbol{y}$ [6].

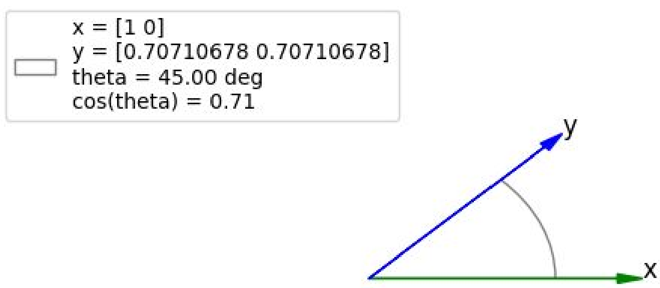

Figure (1) – Exemple simpliste de représentation de 2 vecteurs x et y avec un angle θ de 45 degrés (frobenius-graph-8-2) tracée avec un programme en python (numpy, matplotlib.pyplot)

```
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.patches as patches
...
def draw_angle_arc(theta, radius):
...
def draw_vectors(x, y):
...
```

6. la formule montre que le produit scalaire de deux vecteurs est égal au produit de leurs normes euclidiennes et du cosinus de l'angle θ entre eux. En d'autres termes, le produit scalaire mesure la projection de l'un des vecteurs sur l'autre, normalisée par la norme du vecteur projeté.

```
def create_plot(theta, x, y):
...
x = np.array([1, 0])
y = np.array([np.cos(np.radians(45)), np.sin(np.radians(45))])
theta = np.arccos(np.dot(x, y) / (np.linalg.norm(x) * np.linalg.norm(y)))
create_plot(theta, x, y)
plt.show()
```

6. Fonction Norme

En apprentissage automatique, on mesure généralement la taille des vecteurs à l'aide d'une fonction appelée **norme** et parfois, on a besoin de mesurer la taille d'un vecteur. Formellement, la norme L^p est donnée par :

$$\|\boldsymbol{x}\|_p - \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} \tag{8}$$

pour $p \in \mathbb{R}, p \geq 1$. Les normes, y compris la norme L^p, sont des fonctions qui associent des vecteurs à des valeurs non négatives.
La norme d'un vecteur $\boldsymbol{x}$ mesure la distance de l'origine au point $\boldsymbol{x}$.

Une norme [7] est une fonction f qui satisfait les propriétés suivantes :

1. $f(\boldsymbol{x}) = 0 \Rightarrow \boldsymbol{x} = \boldsymbol{0}$
2. $f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y})$ (propriété de l'inégalité triangulaire)
3. $\forall \alpha \in \mathbb{R}, f(\alpha \boldsymbol{x}) - |\alpha| f(\boldsymbol{x})$ (propriété de l'homogénéité) [8]

Comme le précise (Goodfellow et al., 2016) [2], la norme L^2, avec $p = 2$, est connue sous le nom de norme euclidienne, qui est simplement la distance euclidienne de l'origine au point identifié par $\boldsymbol{x}$. La norme L^2 est utilisée si fréquemment en apprentissage automatique qu'elle est souvent notée simplement $|\boldsymbol{x}|$, avec l'indice 2 omis. Il est également courant de mesurer la taille d'un vecteur en utilisant la norme L^2 au carré, qui peut être calculée simplement comme $\boldsymbol{x}^\top \boldsymbol{x}$.

La norme L^2 au carré est plus pratique à manipuler mathématiquement et pour du calcul sur ordinateur, que la norme L^2 elle-même. Par exemple, chaque dérivée de la norme L^2 au carré par rapport à chaque élément de x ne dépend que de

7. lorsque la norme $f(\boldsymbol{x})$ d'un vecteur $\boldsymbol{x}$ tend vers 0, alors $\boldsymbol{x}$ tend également vers 0. En d'autres termes, cette propriété stipule que si la distance entre un vecteur $\boldsymbol{x}$ et l'origine (0) devient de plus en plus petite (c'est-à-dire que la norme $f(\boldsymbol{x})$ tend vers 0), alors le vecteur $\boldsymbol{x}$ lui-même se rapproche également de l'origine. Cette propriété est importante car elle assure que la norme est une mesure de la "taille" d'un vecteur. Si la norme d'un vecteur est proche de zéro, cela signifie que le vecteur lui-même est petit, ou "proche de zéro".

8. La propriété d'homogénéité d'une fonction f signifie que pour tout scalaire α et tout vecteur $\boldsymbol{x}$ dans le domaine de f, on a$f(\alpha \boldsymbol{x}) = |\alpha| f(\boldsymbol{x})$. Cela signifie que la valeur de la fonction f appliquée à un vecteur multiplié par un scalaire est égale à la valeur absolue de ce scalaire multiplié par la valeur de la fonction appliquée au vecteur initial.

l'élément correspondant de x, tandis que toutes les dérivées de la norme L^2 dépendent de l'ensemble du vecteur. Dans de nombreux contextes, la norme L^2 au carré peut être indésirable car elle augmente très lentement près de l'origine. Dans plusieurs applications d'apprentissage automatique, il est important de faire la distinction entre les éléments qui sont exactement zéro et ceux qui sont petits mais non nuls. Dans ces cas, nous nous tournons vers une fonction qui croît rapidement vers zéro, telle que la norme L^1.

7. Applications en cybersécurité et machine learning

L'article *"Effects of Jacobian Matrix Regularization on the Detectability of Adversarial Samples"* (Eydenberg, 2020) [6] étudie l'utilisation de la régularisation de la *matrice jacobienne*[9] dans la détection d'échantillons adversaires dans les réseaux de neurones. Les échantillons adversaires sont des données manipulées intentionnellement pour induire une erreur dans les modèles de machine learning. Les auteurs de l'article examinent l'effet de la régularisation de la matrice jacobienne sur la capacité d'un réseau de neurones à détecter ces échantillons adversaires. Ils utilisent la **norme de Frobenius** pour mesurer la magnitude des perturbations dans la matrice jacobienne. Ils montrent que la régularisation de la matrice jacobienne peut améliorer la capacité des réseaux de neurones à détecter les échantillons adversaires. Ils discutent également des limites de leur approche et des pistes pour de futures recherches dans ce domaine. En résumé, l'article présente une approche de régularisation de la matrice jacobienne pour améliorer la détection des échantillons adversaires dans les réseaux de neurones, en utilisant la norme de Frobenius pour mesurer la magnitude des perturbations.
Ajouter une pénalité de régularisation à une fonction de perte signifie qu'on ajoute un terme supplémentaire à la fonction de perte originale qui va pénaliser les modèles qui ont des valeurs de paramètres extrêmes. Dans le contexte de leur article, elle prend la forme de la norme de Frobenius de la Jacobienne du modèle. Cette pénalité va encourager le modèle à avoir des valeurs de paramètres plus proches de zéro, ce qui peut aider à éviter le surapprentissage et à améliorer la capacité de généralisation du modèle.
Les résultats suggèrent que cette approche peut être prometteuse pour renforcer la sécurité des systèmes de machine learning contre les attaques adversaires.

8. Applications en cybersécurité et cryptographie

L'article *"Secure Control of Nonlinear Systems Using Semi-Homomorphic Encryption"* de (Lin et al, 2021) [7] propose une approche pour garantir la confidentialité et la sécurité des données dans les systèmes de contrôle non linéaires. Les auteurs utilisent une méthode de cryptage semi-homomorphe pour chiffrer les données et garantir leur sécurité. La méthode consiste à ajouter du bruit aléatoire au signal de commande pour prévenir les attaques par analyse de trafic et à ajouter une fonction de

9. Matrice des dérivées partielles d'une fonction vectorielle

perte qui combine la **norme de Frobenius** et la norme L_2 pour régulariser le processus d'optimisation. Les auteurs ont testé leur méthode sur un système de contrôle de drone et ont montré que leur approche était efficace pour garantir la sécurité et la confidentialité des données.

La norme de Frobenius est utilisée dans l'article comme une mesure de la complexité du modèle, ce qui permet de régulariser le processus d'optimisation et d'éviter le surapprentissage. En ajoutant une pénalité pour la norme de Frobenius de la matrice Jacobienne dans la fonction de perte, les auteurs ont pu garantir la stabilité numérique du processus d'optimisation et éviter les explosions de gradients. La norme L_2 a également été utilisée pour contrôler la taille des données de bruit ajoutées au signal de commande. La combinaison de ces deux normes a permis de régulariser le processus d'optimisation et de garantir la sécurité et la confidentialité des données dans le système de contrôle non linéaire.

Spécifiquement, la ligne 3 de l'algorithme proposé par (Lin et al, 2021) [7] (ligne spécifique reprise plus bas) consiste à calculer une matrice $\bar{K}$ qui minimise la distance entre A et K, où K est une matrice connue et A est une matrice de décision. Cette distance est mesurée en utilisant la norme de Frobenius, c'est-à-dire la racine carrée de la somme des carrés des éléments de la différence entre A et K. Ainsi, l'objectif est de trouver la matrice $\bar{K}$ qui est la plus proche de K en termes de **la norme de Frobenius**. Cela permet de garantir que la matrice de décision obtenue est la plus proche possible de la matrice K tout en préservant la sécurité et la confidentialité des mesures :

Algorithm 1 Secure and private implementation of the static controller with encrypted measurements

...

3. Compute $\bar{K} \leftarrow \arg\min_{A \in \mathbb{Q}(n_1,m_1)^{n_u \times n_y}} \|A - K\|_F$

...

Références

[1] Linear Algebra : norms.
URL https://www.deeplearningbook.org/contents/linear_algebra.html

[2] Ian Goodfellow and Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016, http://www.deeplearningbook.org.

[3] Biographie de Ferdinand Frobenius.
URL https://www.bibmath.net/bios/index.php?action=affiche&quoi=frobenius

[4] F. G. Frobenius, G. L. Frobenius, Über Matrizen aus positiven Elementen. II., no. v. 2, Königliche Akademie der Wissenschaften, 1909.
URL https://books.google.ca/books?id=7I8TnQEACAAJ

[5] Norme matricielle.
URL https://fr.wikipedia.org/wiki/Norme_matricielle

[6] Eydenberg, Michael Shannon and Khanna, Kanad and Custer, Ryan, Effects of Jacobian Matrix Regularization on the Detectability of Adversarial Samples (9 2020). doi:10.2172/1763568.
URL https://www.osti.gov/biblio/1763568

[7] Lin, Yankai and Zhang, Jie and Chen, Weiqiang and Sun, Jing and Xu, Xiaowei, Secure control of nonlinear systems using semi-homomorphic encryption, IEEE Transactions on Control of Network Systems 8 (3) (2021) 1355–1366.
URL https://people.eng.unimelb.edu.au/dnesic/archive-hall/yankai-etal-CDC2018.pdf