Feuille de théorie 4

Formes linéaires

Definition 1. *Soient $a_1, \dots, a_n \in \mathbb{R}$ des nombres. Une* <u>*forme linéaire*</u> *associée à $(a_1, \dots, a_n)$ est une fonction $f : \mathbb{R}^n \to \mathbb{R}$ donnée par*

$$f(x_1, \dots, x_n) = a_1x_1 + a_2x_2 + \dots + a_nx_n = \sum_{k=1}^{n} a_kx_k.$$

Si on note $A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ *et* $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, *alors f peut s'écrire comme*

$$f(\boldsymbol{x}) = A^\top \boldsymbol{x}.$$

Par exemple,

$$f(x_1, x_2) = \begin{pmatrix} 3 \\ 4 \end{pmatrix}^\top \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 + 4x_2.$$

Remark 1. *Les vecteurs colonnes = matrices $n \times 1$ sont généralement notés soit $\boldsymbol{x}$ (en gras) soit $\vec{x}$ (avec une flèche au-dessus). Les deux notations sont standard. La deuxième est plus fréquemment utilisée à la main, alors que la seconde est plus standard dans les publications.*

Voici une autre définition :

Definition 2. *Une forme linéaire $f : \mathbb{R}^n \to \mathbb{R}$ est une fonction qui satisfait*

$$f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$$

pour tous $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

Les deux définitions sont <u>équivalentes</u> :

Theorem 1. *Si $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ pour tous $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, alors il existe des nombres $a_1, \dots, a_n$ tels que $f(\boldsymbol{x}) = a_1x_1 + \dots + a_nx_n$.*

Le vecteur colonne A^T a un nom spécial :

Definition 3. *Si $f : \mathbb{R}^n \to \mathbb{R}$ est une forme linéaire $f(\boldsymbol{x}) = A^\top \boldsymbol{x}$, alors $A^\top$ est appelé le* <u>*gradient*</u> *de f.*

Plus tard, nous introduirons les gradients d'autres fonctions.

Definition 4. *Un graphe d'une fonction* $f : \mathbb{R}^n \to \mathbb{R}$ *est l'ensemble*

$$\left\{(x_1, \dots, x_n, y) \in \mathbb{R}^{n+1} : y = f(x_1, \dots, x_n)\right\}.$$

Definition 5 (Ligne, plan, hyperplan)**.**

- *Une ligne passant par l'origine est le graphe d'une forme linéaire* $f : \mathbb{R} \to \mathbb{R}$ *.*
- *Un plan passant par l'origine est le graphe d'une forme linéaire* $f : \mathbb{R}^2 \to \mathbb{R}$ *.*
- *Un hyperplan passant par l'origine est le graphe d'une forme linéaire* $f : \mathbb{R}^n \to \mathbb{R}$ *.*

Vérifiez que cela correspond à votre compréhension intuitive de ce que sont une ligne et un plan en dessinant quelques figures.

Definition 6. *Si* $f(\boldsymbol{x}) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$ *, alors* a_i *est considéré comme la pente de* f *dans la direction* x_i *.*

Courbes/ensembles de niveau de formes linéaires

Definition 7. *L'ensemble de niveau d'une fonction* $f : \mathbb{R}^n \to \mathbb{R}$ *avec la valeur* $c \in \mathbb{R}$ *est un ensemble où cette fonction prend la valeur* c *:*

$$\{(x_1, \dots, x_n) : f(x_1, \dots, x_n) = c\}.$$

Deux remarques :

- Les ensembles de niveau des formes linéaires sont des lignes, des plans ou des hyperplans ne passant pas nécessairement par zéro.
- Le gradient est orthogonal à l'ensemble de niveau.

Transformation linéaire

Definition 8. *Soit* $A \in M_{m,n}$ *une matrice. Une transformation linéaire associée à* A *est une fonction* $f : \mathbb{R}^n \to \mathbb{R}^m$ *donnée par*

$$f(\boldsymbol{x}) = A\boldsymbol{x}.$$

Plus explicitement,

$$f(x_1, \dots, x_n) = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

où les nombres y_i *sont donnés par*

$$y_i = (A\boldsymbol{x})_i = \sum_{k=1}^{n} a_{ik}x_k.$$

Comme pour les formes linéaires, il existe une autre définition :

Definition 9. *Une transformation linéaire est une fonction* $f : \mathbb{R}^n \to \mathbb{R}^m$ *qui satisfait*

$$f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$$

pour tous $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ *.*

Et comme pour les formes linéaires, cette définition est équivalente à la première :

Theorem 2. *Si* $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ *pour tous* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ *, alors il existe une matrice* $A \in M_{m,n}$ *telle que* $f(\boldsymbol{x}) = A\boldsymbol{x}$ *.*

Proof. Puisque $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ pour tous $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, nous avons

$$f(\boldsymbol{x}) = f\left(\sum_{j=1}^{n} x_j \boldsymbol{e}_j\right) = \sum_{j=1}^{n} x_j f(\boldsymbol{e}_j),$$

où $\boldsymbol{e}_j$ est le j^{th} vecteur de la base standard :

$$\boldsymbol{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow i \\ \\ \\ \end{matrix}$$

Par conséquent, la i^{th} composante de $\boldsymbol{y} = f(\boldsymbol{x})$ est donnée par

$$y_i = \sum_{j=1}^{n} x_j (f(\boldsymbol{e}_j))_i = \sum_{j=1}^{n} A_{ij} x_j,$$

où $A_{ij} = (f(\boldsymbol{e}_j))_i$. □

Voici quelques remarques :

- Notez la position de m et n : si $A \in M_{m,n}$, alors $f(\boldsymbol{x}) = A\boldsymbol{x}$ est une fonction $\mathbb{R}^n \to \mathbb{R}^m$.
- En d'autres termes, m est la dimension de sortie et n est la dimension d'entrée.
- La définition mystérieuse du produit matriciel vient en fait du fait suivant : si $f : \mathbb{R}^n \to \mathbb{R}^p$ et $g : \mathbb{R}^p\ to \mathbb{R}^m$ satisfont

  $$f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}) \quad \text{et} \quad g(\boldsymbol{x}'+\boldsymbol{y}') = g(\boldsymbol{x}') + g(\boldsymbol{y}')$$

  pour tous $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ et tous $\boldsymbol{x}', \boldsymbol{y}' \in \mathbb{R}^p$, alors leur composition

  $$h : \mathbb{R}^n \to \mathbb{R}^m : h(\boldsymbol{x}) = g(f(\boldsymbol{x}))$$

  satisfait également cette propriété :

  $$h(\boldsymbol{x}+\boldsymbol{y}) = h(\boldsymbol{x}) + h(\boldsymbol{y})$$

  pour tous $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Nous savons d'après le théorème ci-dessus que $h(\boldsymbol{x}) = C\boldsymbol{x}$ avec un certain C , mais est-il possible de trouver C ?

- Si f est représentée par la matrice B (c'est-à-dire $f(\boldsymbol{x}) = B\boldsymbol{x}$) et g est représentée par la matrice A (c'est-à-dire $g(\boldsymbol{x}') = A\boldsymbol{x}'$), alors h est représentée par $C = AB$.
- En d'autres termes, le produit des matrices représente la composition des transformations linéaires.
- Comparer:

$$A \in M_{m,p}, \;\; B \in M_{p,n} \implies AB \in M_{m,n}$$
$$f : \mathbb{R}^n \to \mathbb{R}^p, \;\; g : \mathbb{R}^p \to \mathbb{R}^m \implies g(f(\boldsymbol{x})) : \mathbb{R}^n \to \mathbb{R}^m.$$

Exemples de transformations linéaires

Rotations

Si $\theta \in [0, 2\pi]$ est un angle, la transformation linéaire $r_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ associée à la matrice R_θ par $r_\theta(\boldsymbol{x}) = R_\theta \boldsymbol{x}$, où

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

fait pivoter $\boldsymbol{x}$ de l'angle θ autour de $\mathbf{0}$.

Par exemple, si $\theta = -\frac{\pi}{3}$, nous avons

$$R_{-\pi/3} = \begin{pmatrix} \cos\left(-\frac{\pi}{3}\right) & -\sin\left(-\frac{\pi}{3}\right) \\ \sin\left(-\frac{\pi}{3}\right) & \cos\left(-\frac{\pi}{3}\right) \end{pmatrix} = \begin{pmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{pmatrix}.$$

Dilatation/homothétie/mise à l'échelle

Si $k \in \mathbb{R}$, la transformation linéaire $h : \mathbb{R}^2 \to \mathbb{R}^2$ donnée par

$$h(\boldsymbol{x}) = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

est la dilatation ou homothétie ou mise à l'échelle par un facteur de k .

- Si $|k| > 1$, h augmente les tailles
- Si $|k| < 1$, h diminue les tailles
- Si $k > 0$, h conserve les directions
- Si $k < 0$, h inverse les directions

Résolution de systèmes linéaires

Étant donné $\boldsymbol{y} \in \mathbb{R}^m$ et $A \in M_{m,n}$, nous voulons résoudre $A\boldsymbol{x} = \boldsymbol{y}$ pour $\boldsymbol{x}$. Il y a deux questions naturelles :

- Est-ce possible ? En d'autres termes, existe-t-il $\boldsymbol{x}$ tel que $A\boldsymbol{x} = \boldsymbol{y}$?
- Est-ce que $\boldsymbol{x}$ est unique ? En d'autres termes, est-il possible que $A\boldsymbol{x} = \boldsymbol{y} = A\boldsymbol{x}'$ avec $\boldsymbol{x} \neq \boldsymbol{x}'$?

Remark 2. *Dans le cours précédent, nous avons vu que* $A\boldsymbol{x} = \boldsymbol{y}$ *peut être résolu par* $\boldsymbol{x} = A^{-1}\boldsymbol{y}$, *mais seulement si* A^{-1} *existe. Rappelons que* A^{-1} *n'a de sens que pour les matrices* <u>*carrées*</u>*. Nous sommes maintenant intéressés à résoudre* $A\boldsymbol{x} = \boldsymbol{y}$ <u>*sans supposer que la matrice est carrée*</u>*.*

Le théorème suivant donne la réponse (sans preuve) :

Theorem 3.

- *Si le rang de* A *coïncide avec la* <u>*dimension de sortie*</u>*, c'est-à-dire* $\operatorname{Rank} A = m$, *alors* $\boldsymbol{x}$ *existe*
- *Si le rang de* A *est* <u>*inférieur à la dimension de sortie*</u>*, c'est-à-dire* $\operatorname{Rank} A < m$, *alors* ***il existe*** *certains* $\boldsymbol{y} \in \mathbb{R}^m$ *tels que l'équation* $A\boldsymbol{x} = \boldsymbol{y}$ *est insoluble*
- *Si le rang de* A *coïncide avec la* <u>*dimension d'entrée*</u>*, c'est-à-dire* $\operatorname{Rank} A = n$, *alors la solution est unique (si elle existe ; voir les points précédents)*
- *Si le rang de* A *est* <u>*inférieur à la dimension d'entrée*</u>*, alors la solution n'est pas unique (si elle existe ; voir les points précédents)*

En particulier,

> Si le rang de A coïncide avec <u>les dimensions d'entrée et de sortie</u>
>
> $$\operatorname{Rank} A = m = n,$$
>
> la solution existe et est unique.

Voici quelques remarques :

- Il y a <u>trois possibilités</u> soit (1) il y a une solution unique <u>ou</u> (2) il n'y a pas de solution <u>ou</u> (3) il y a <u>un nombre infini</u> de solutions.
- Exemple du cas (1) :

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 0 \end{cases} \implies x_1 = x_2 = 1.$$

- Exemple du cas (2) :

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases} \implies \text{pas de solution}$$

- Exemple du cas (3) :

$$\begin{cases} x_1 + x_2 = 1 \\ 2x_1 + 2x_2 = 2 \end{cases} \implies x_2 = 1 - x_1, \quad x_2 \text{ n'importe quel nombre.}$$

- Si $\operatorname{Rank} A < m$ (rang inférieur à la dimension de sortie), $A\boldsymbol{x} = \boldsymbol{y}$ peut être soluble pour certains $\boldsymbol{y}$, mais pas pour tous les $\boldsymbol{y}$.
- Si $\operatorname{Rank} A < n$ (rang inférieur à la dimension d'entrée), il peut y avoir de nombreuses solutions et nous voulons généralement décrire toute la famille des solutions.

Résolution de systèmes linéaires à l'aide de transformations élémentaires

Afin de résoudre un système linéaire

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & y_1 \\ a_{21}x_2 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & y_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_m & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & y_m \end{array}\right.$$

nous réécrivons ce système d'équations sous forme matricielle

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

et utilisons <u>la même idée</u> que nous avons utilisée pour trouver le rang, le déterminant et l'inverse : appliquer des transformations élémentaires <u>aux deux côtés de cette équation</u> (c'est-à-dire à A et à $\boldsymbol{y}$).

Le <u>but</u> est le même qu'avec la recherche de l'inverse : transformer la matrice à gauche en la matrice identité I.

<u>Cependant</u>, cela peut ne pas être possible si A n'est pas inversible. Si $\operatorname{Rank} A < m$, à un moment donné, nous rencontrons des <u>lignes nulles</u> :

$$i \rightarrow \begin{pmatrix} * & * & * & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} * \\ \vdots \\ \tilde{y}_i \\ \vdots \\ * \end{pmatrix} \leftarrow i$$

- Si la i^{th} valeur $\tilde{y}_i$ n'est pas nulle, aucun choix de $\boldsymbol{x}$ ne produira une solution. Dans ce cas, nous concluons qu'il n'y a pas de solution.
- Si $\tilde{y}_i = 0$, nous pouvons simplement jeter la ligne nulle et continuer à résoudre le problème.

Si, d'autre part, $\operatorname{Rank} A < n$, à un moment donné, nous pouvons rencontrer des colonnes nulles :

$$\begin{array}{c}\left(\begin{array}{ccccc} * & \dots & 0 & \dots & * \\ \vdots & \vdots & \vdots & \dots & \vdots \\ * & \dots & 0 & \dots & * \end{array}\right) \\ \begin{array}{ccccc} & & \uparrow & & \\ & & j & & \end{array}\end{array} \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} * \\ \vdots \\ * \end{pmatrix}.$$

- Si la j^{th} colonne est nulle, chaque choix de $x_j \in \mathbb{R}$ donne une solution.
- Par conséquent, la solution (si elle existe) n'est pas unique.

Remark 3. *En pratique, nous n'avons fréquemment pas besoin de simplifier la matrice jusqu'à la fin. Prenons un exemple. Si nous avons réduit le système à la forme*

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix},$$

nous voyons immédiatement que le problème n'a pas de solutions car il y a une valeur non nulle 2 *contre une ligne nulle. Si, d'autre part,*

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 0 \end{pmatrix},$$

alors nous pouvons, sans simplifier davantage la matrice, revenir à la forme "système d'équations"

$$\begin{cases} x_1 + x_2 + x_3 = 2 \\ x_2 + 3x_3 = 6 \end{cases}$$

et le résoudre à la main :

$$x_2 = 6 - 3x_3, \quad x_1 = 2 - x_3 - (6 - 3x_3) = -4 + 2x_3,$$

où x_3 *est un nombre quelconque.*

Comme nous l'avons fait pour trouver l'inverse, il est utile d'introduire la notation matrice augmentée pour résoudre les systèmes linéaires :

Definition 10. *Associons à un système d'équations linéaires*

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & y_1 \\ a_{21}x_2 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & y_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_m & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & y_m \end{array}\right.$$

une <u>*matrice augmentée*</u> $(A \mid \boldsymbol{y})$ *écrite comme*

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & y_1 \\ a_{21} & a_{22} & \dots & a_{2n} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & y_m \end{array}\right).$$

De même à ce que nous avons fait pour trouver l'inverse, nous pouvons maintenant résoudre ce système d'équations linéaires en appliquant des transformations élémentaires à la matrice augmentée $(A \mid \boldsymbol{y})$.