

# DM n°01 — Classements et cycles cachés

Algèbre linéaire · Espaces  
euclidiens · Matrices · Projections  
orthogonales

Comparer des objets deux à deux ne conduit pas toujours à un classement cohérent : on peut préférer  $A$  à  $B$ ,  $B$  à  $C$ , puis  $C$  à  $A$ . On cherche ici à distinguer ce qui peut être expliqué par des scores globaux de ce qui subsiste sous forme de cycles.

On appelle graphe fini simple non orienté un couple  $G = (V, E)$  où

$$V = \{1, \dots, n\}$$

et où  $E$  est un ensemble de paires  $\{i, j\}$  de sommets distincts. Deux sommets  $i$  et  $j$  sont adjacents lorsque  $\{i, j\} \in E$ . Une arête  $\{i, j\}$  est dite incidente aux sommets  $i$  et  $j$ . Le degré d'un sommet est le nombre d'arêtes qui lui sont incidentes.

Un chemin est une suite de sommets  $(v_0, \dots, v_k)$  telle que  $\{v_{r-1}, v_r\} \in E$  pour tout  $r$ . Le graphe est connexe lorsque deux sommets quelconques peuvent être reliés par un chemin. Un cycle est un chemin fermé  $(v_0, \dots, v_k)$  avec  $k \geq 3$ , dans lequel  $v_0, \dots, v_{k-1}$  sont distincts et  $v_k = v_0$ . Un cycle de longueur 3 est appelé un triangle.

On considère désormais un graphe fini simple connexe  $G = (V, E)$  à  $n$  sommets et  $m$  arêtes. On fixe un ordre des arêtes. Pour chaque arête  $\{i, j\}$ , on choisit arbitrairement un sens, noté  $i \rightarrow j$  ou  $j \rightarrow i$  ; ce choix ne fait que fixer des signes.

On note  $B \in \mathcal{M}_{n,m}(\mathbb{R})$  la matrice d'incidence définie ainsi : si l'arête  $e$  est orientée de  $i$  vers  $j$ , sa colonne dans  $B$  contient  $-1$  en ligne  $i$ ,  $+1$  en ligne  $j$  et  $0$  ailleurs. Les coordonnées d'un vecteur de  $\mathbb{R}^n$  sont indexées par les sommets et celles d'un vecteur de  $\mathbb{R}^m$  par les arêtes.

## I Incidence et orthogonalité

- 1 Soit  $e$  une arête orientée de  $i$  vers  $j$ .
  - (a) Pour  $p \in \mathbb{R}^n$ , exprimer la coordonnée  $(B^\top p)_e$  en fonction de  $p_i$  et  $p_j$ .
  - (b) Pour  $x \in \mathbb{R}^m$ , exprimer  $(Bx)_i$  à l'aide des coordonnées de  $x$  portées par les arêtes incidentes au sommet  $i$ .
- 2 Déterminer  $\ker(B^\top)$ . En déduire le rang de  $B$  puis la dimension de  $\ker B$ .
- 3 Établir la décomposition orthogonale

$$\mathbb{R}^m = \text{Im}(B^\top) \oplus^\perp \ker B.$$

En déduire que, pour tout  $x \in \mathbb{R}^m$ , il existe un unique couple  $(g, c)$  tel que

$$x = g + c, \quad g \in \text{Im}(B^\top), \quad c \in \ker B.$$

Si  $g = B^\top p$ , caractériser tous les vecteurs  $p$  qui donnent le même vecteur  $g$ .

## II Duels et meilleur score global

On dispose maintenant d'un vecteur  $y \in \mathbb{R}^m$  de comparaisons deux à deux. Si l'arête  $e$  est orientée de  $i$  vers  $j$ , la quantité  $y_e > 0$  signifie que la comparaison avantage  $j$  par rapport à  $i$  ; la valeur absolue  $|y_e|$  mesure l'intensité de cette comparaison.

Un système de scores  $p \in \mathbb{R}^n$  prédit sur l'arête  $i \rightarrow j$  la différence  $p_j - p_i$ , c'est-à-dire la coordonnée correspondante de  $B^\top p$ . On cherche donc à rendre

$$\|y - B^\top p\|^2$$

aussi petite que possible. Ajouter une même constante à tous les scores ne change aucune différence  $p_j - p_i$  ; la condition  $p_1 + \dots + p_n = 0$  permettra de fixer cette liberté.

#### 4 Caractériser tous les minimiseurs de l'application

$$p \mapsto \|y - B^\top p\|^2.$$

En déduire qu'il existe un unique minimiseur vérifiant

$$p_1 + \dots + p_n = 0.$$

#### 5 On pose

$$L = BB^\top.$$

Montrer que  $L$  est symétrique et que  $p^\top L p \geq 0$  pour tout  $p \in \mathbb{R}^n$ . Déterminer  $\ker L$  et exprimer les coefficients de  $L$  directement à partir du graphe.

#### 6 On inverse l'orientation d'une seule arête $e$ et, pour représenter les mêmes données, on remplace simultanément $y_e$ par $-y_e$ . Montrer que le minimiseur de somme nulle et la valeur minimale de $\|y - B^\top p\|$ ne dépendent pas du choix initial des orientations.

### III Un exemple complet à quatre objets

On appelle graphe complet à quatre sommets, noté  $K_4$ , le graphe dont toute paire de sommets distincts forme une arête. On oriente chaque arête du plus petit indice vers le plus grand. Dans l'ordre

$$12, 13, 14, 23, 24, 34,$$

on observe

$$y = (3, 3, 6, 3, 4, 2)^\top.$$

#### 7 Construire la matrice $B$ , puis déterminer le score $p$ de somme nulle qui minimise $\|y - B^\top p\|^2$ . Calculer l'écart

$$c = y - B^\top p$$

ainsi que le rapport

$$\frac{\|c\|^2}{\|y\|^2}.$$

Décrire les arêtes sur lesquelles  $c$  est non nul et interpréter les signes obtenus dans le contexte des comparaisons.

## IV Triangles et cycles qui leur échappent

On revient au graphe général  $G$ . On sélectionne  $q$  triangles  $T_1, \dots, T_q$  et, pour chacun d'eux, on choisit un sens de parcours.

À chaque triangle  $T_k$ , on associe un vecteur  $c_k \in \mathbb{R}^m$  : sur une arête du triangle il vaut  $+1$  si le sens de parcours coïncide avec l'orientation choisie de l'arête,  $-1$  dans le cas contraire, et il vaut  $0$  sur les autres arêtes. On rassemble ces vecteurs en colonnes dans

$$C = (c_1 \quad \cdots \quad c_q) \in \mathcal{M}_{m,q}(\mathbb{R}).$$

- 8 Montrer que  $BC = 0$ . En posant

$$H = \ker B \cap \ker(C^\top),$$

établir la décomposition orthogonale

$$\mathbb{R}^m = \operatorname{Im}(B^\top) \oplus^\perp \operatorname{Im} C \oplus^\perp H.$$

En déduire une expression de  $\dim H$  en fonction de  $m$ ,  $n$  et  $\operatorname{rg} C$ .

- 9 Revenons à  $K_4$ . On sélectionne les quatre triangles 123, 124, 134, 234. Pour un triangle  $ijk$  avec  $i < j < k$ , on choisit le sens

$$i \rightarrow j \rightarrow k \rightarrow i.$$

Déterminer  $\operatorname{rg} C$  et l'espace  $H$ . Situer le vecteur  $c$  de la question 7 dans la décomposition de la question 8.

- 10 Considérons le graphe formé uniquement par les quatre arêtes du cycle

$$1 - 2 - 3 - 4 - 1,$$

sans diagonale. On les ordonne 12, 23, 34, 14 et on les oriente du plus petit indice vers le plus grand. Aucun triangle n'étant sélectionné, on prend  $\operatorname{Im} C = \{0\}$ .

Déterminer  $H$  et en exhiber un générateur explicite. Comparer ce résultat avec celui de la question 9.

- 11 Pour un vecteur de comparaisons quelconque  $y \in \mathbb{R}^m$ , montrer qu'il existe une unique décomposition

$$y = g + t + h,$$

avec

$$g \in \operatorname{Im}(B^\top), \quad t \in \operatorname{Im} C, \quad h \in H.$$

Si l'on impose  $g = B^\top p$  avec  $p_1 + \cdots + p_n = 0$ , que peut-on dire de  $p$  ? Si l'on écrit  $t = C\alpha$  avec  $\alpha \in \mathbb{R}^q$ , à quelle condition sur  $C$  le vecteur  $\alpha$  est-il unique ?

Interpréter finalement les trois composantes dans le problème de classement à partir de comparaisons deux à deux.