

Plaquette 5 – Synthèse : réseaux, PageRank et modèles aléatoires

Graphes et marches aléatoires — Terminale maths expertes

Corrections détaillées

Boîte à outils

Toute la boîte à outils du chapitre.

- **Matrice d'adjacence** M : $m_{ij} = 1$ s'il existe une arête (ou un arc) de i vers j . Le coefficient $(i; j)$ de M^n compte les **chemins de longueur** n de i vers j .
- **Degrés** : somme des lignes ; la somme des degrés vaut deux fois le nombre d'arêtes (graphe non orienté).
- **Connexité** : tous les coefficients de $I + M + \dots + M^{p-1}$ sont strictement positifs, pour un graphe d'ordre p .
- **Marche aléatoire simple** : $t_{ij} = \frac{1}{\deg(j)}$ si i et j sont voisins. Distribution invariante : $p_i = \frac{\deg(i)}{2 \times \text{nombre d'arêtes}}$.
- **Graphe probabiliste** : $P_n = T^n P_0$; la distribution invariante vérifie $TP = P$ avec somme 1.
- **Cas à deux états** : $p_{n+1} = (1 - a - b)p_n + b$, limite $\frac{b}{a+b}$, convergence dès que $|1 - a - b| < 1$.
- **PageRank** : on modélise un internaute qui suit au hasard les liens d'une page ; la distribution invariante mesure l'**importance** de chaque page.
- **Le dernier geste, toujours** : traduire le résultat numérique en une phrase sur le réseau étudié.

Correction de l'exercice 1 – Un réseau de métro

a) On compte les voisins de chaque station :

Station	S_1	S_2	S_3	S_4	S_5
Degré	2	3	4	3	2

- Somme des degrés : $2 + 3 + 4 + 3 + 2 = 14$, donc $\frac{14}{2} = 7$ liaisons ✓ (on les recompte dans l'énoncé : il y en a bien sept).

b) La distribution invariante d'une marche aléatoire simple est proportionnelle aux degrés :

$$p_i = \frac{\deg(i)}{2 \times 7} = \frac{\deg(i)}{14}$$

Station	S_1	S_2	S_3	S_4	S_5
Proportion	$\frac{2}{14}$	$\frac{3}{14}$	$\frac{4}{14}$	$\frac{3}{14}$	$\frac{2}{14}$
Environ	14,3 %	21,4 %	28,6 %	21,4 %	14,3 %

- **Contrôle** : la somme vaut $\frac{14}{14} = 1$ ✓.
- **Interprétation** : la station S_3 , la mieux connectée, concentre près de 29 % du trafic. C'est exactement le raisonnement qui justifie de dimensionner une station de

correspondance plus grande que les autres.

Réponse. a) degrés 2, 3, 4, 3, 2 et 7 liaisons · b) proportions $\frac{\deg(i)}{14}$, soit 28,6 % pour S_3 .

Correction de l'exercice 2 – Combien d'itinéraires ?

a) Le coefficient $(1; 4)$ de M^2 vaut 2 : il y a **deux** itinéraires de longueur 2.

— Ce sont $1 - 2 - 4$ et $1 - 3 - 4$.

— **Vérification** : les voisins de la ville 1 sont 2 et 3, et toutes deux sont reliées à 4 ✓.

b) Le coefficient $(1; 4)$ de M vaut 0 : il n'y a **aucune** liaison directe entre les villes 1 et 4.

— **Lecture d'ensemble** : ces deux villes sont pourtant reliées — en deux étapes, et de deux façons différentes. Le réseau est donc **robuste** entre 1 et 4 : la panne d'une correspondance laisse l'autre itinéraire disponible.

— **Remarque sur la matrice** : M^2 a ses coefficients diagonaux égaux aux degrés (2, 3, 3, 2) ✓, ce qui donne 5 arêtes au total.

Réponse. a) deux : $1 - 2 - 4$ et $1 - 3 - 4$ · b) non, le coefficient de M est nul.

Correction de l'exercice 3 – Le PageRank d'un mini-web

a) Chaque **colonne** décrit les départs depuis une page, les liens étant équiprobables.

$$T = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/3 & 0 & 1/2 & 1/2 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 0 & 0 \end{pmatrix}$$

— Colonne 1 : trois liens, donc $\frac{1}{3}$ chacun vers les pages 2, 3, 4 ✓.

— Colonnes de somme 1 ✓.

b) On calcule TP avec $P = \begin{pmatrix} 0,3 \\ 0,3 \\ 0,3 \\ 0,1 \end{pmatrix}$:

— Ligne 1 : $\frac{1}{2} \times 0,3 + \frac{1}{2} \times 0,3 = 0,15 + 0,15 = 0,3$ ✓.

— Ligne 2 : $\frac{1}{3} \times 0,3 + \frac{1}{2} \times 0,3 + \frac{1}{2} \times 0,1 = 0,1 + 0,15 + 0,05 = 0,3$ ✓.

— Ligne 3 : $\frac{1}{3} \times 0,3 + \frac{1}{2} \times 0,3 + \frac{1}{2} \times 0,1 = 0,3$ ✓.

— Ligne 4 : $\frac{1}{3} \times 0,3 = 0,1$ ✓.

Conclusion. $TP = P$ et la somme vaut 1 : c'est bien la distribution invariante.

— **Classement** : les pages 1, 2 et 3 sont à égalité (30 % chacune), loin devant la page 4 (10 %).

— **Pourquoi la page 4 est faible** : une seule page pointe vers elle (la page 1), et encore en partageant son « vote » entre trois destinations. C'est le principe du PageRank : ce n'est pas le **nombre** de liens entrants qui compte, mais l'importance des pages qui les émettent, divisée par leur nombre de liens sortants.

Réponse. b) $TP = P$ ✓ : les pages 1, 2, 3 à 30 %, la page 4 à 10 %.

Correction de l'exercice 4 – Fiabilité d'un serveur

a) Ici $a = 0,12$ (probabilité de quitter l'état disponible) et $b = 0,48$.

— Distribution invariante : $\frac{b}{a+b} = \frac{0,48}{0,60} = 0,8$.

— **Le taux de disponibilité est de 80 %** : le serveur est saturé une heure sur cinq en moyenne.

b) La relation est $p_{n+1} = (1 - 0,12 - 0,48)p_n + 0,48 = 0,4p_n + 0,48$, avec $p_0 = 1$.

— $p_n - 0,8$ est géométrique de raison 0,4, avec $p_0 - 0,8 = 0,2$:

$$p_n = 0,8 + 0,2 \times 0,4^n$$

— **Vérification** : $p_1 = 0,8 + 0,08 = 0,88$, et le direct donne $0,4 + 0,48 = 0,88$ ✓.

Seuil. $0,8 + 0,2 \times 0,4^n < 0,82$ équivaut à $0,4^n < 0,1$.

— $n > \frac{\ln 0,1}{\ln 0,4} \approx \frac{-2,303}{-0,916} \approx 2,51$.

— **Conclusion** : dès la troisième heure. Vérification : $p_2 = 0,832$ et $p_3 = 0,8128 < 0,82$ ✓.

Réponse. a) 80 % · b) $p_n = 0,8 + 0,2 \times 0,4^n$; seuil franchi à $n = 3$.

Correction de l'exercice 5 - Détecter un triangle

a) Chaque triangle est compté 6 fois dans la trace de R^3 : trois sommets de départ possibles, deux sens de parcours.

— Nombre de triangles : $\frac{18}{6} = 3$.

— **Interprétation** : il existe trois groupes de trois personnes toutes amies entre elles.

b) Le coefficient diagonal de R^2 au sommet i vaut $\deg(i)$: la trace de R^2 est donc la **somme des degrés**.

— Or la somme des degrés vaut deux fois le nombre d'arêtes : $2 \times 7 = 14$.

— **La trace de R^2 vaut donc 14.**

— **Ce que ces deux nombres mesurent** : la trace de R^2 donne le volume de relations, celle de R^3 la densité de « triangles » — un réseau où beaucoup d'amis de mes amis sont mes amis. Les sociologues appellent ce rapport le **coefficient de clustering**.

Réponse. a) 3 triangles · b) 14.

Correction de l'exercice 6 - Une épidémie à trois états

a) $P_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, donc P_1 est la **première colonne** de T :

$$P_1 = \begin{pmatrix} 0,6 \\ 0,2 \\ 0,2 \end{pmatrix}$$

— $P_2 = TP_1$:

— Ligne S : $0,6 \times 0,6 + 0,2 \times 0,2 + 0,2 \times 0,2 = 0,36 + 0,04 + 0,04 = 0,44$.

— Ligne M : $0,2 \times 0,6 + 0,7 \times 0,2 + 0,2 \times 0,2 = 0,12 + 0,14 + 0,04 = 0,30$.

— Ligne I : $0,2 \times 0,6 + 0,1 \times 0,2 + 0,6 \times 0,2 = 0,12 + 0,02 + 0,12 = 0,26$.

— $P_2 = (0,44 ; 0,30 ; 0,26)$, de somme 1 ✓.

b) On calcule TP avec $P = \left(\frac{1}{3}; 0,4; \frac{4}{15}\right)$:

- Ligne S :
 $0,6 \times \frac{1}{3} + 0,2 \times 0,4 + 0,2 \times \frac{4}{15} = 0,2 + 0,08 + \frac{0,8}{15} \approx 0,2 + 0,08 + 0,0533 = 0,3333 \checkmark$.
 - Ligne M : $0,2 \times \frac{1}{3} + 0,7 \times 0,4 + 0,2 \times \frac{4}{15} \approx 0,0667 + 0,28 + 0,0533 = 0,40 \checkmark$.
 - Ligne I : $0,2 \times \frac{1}{3} + 0,1 \times 0,4 + 0,6 \times \frac{4}{15} \approx 0,0667 + 0,04 + 0,16 = 0,2667 \checkmark$.
 - Somme : $\frac{1}{3} + 0,4 + \frac{4}{15} = \frac{5+6+4}{15} = 1 \checkmark$.
 - **Conclusion sanitaire** : à long terme, 40 % de la population est malade — un régime endémique qui se maintient indéfiniment.
- Réponse.** a) $(0,6; 0,2; 0,2)$ puis $(0,44; 0,30; 0,26)$ · b) $TP = P$ et somme 1 $\checkmark$.

Correction de l'exercice 7 - Réseau vulnérable

- a) **Oui** : de proche en proche, on va de A à E par le chemin $A - B - C - D - E$.
- Toutes les paires de serveurs sont reliées, le plus long chemin étant de longueur 4.
 - C'est cohérent avec le critère du cours : pour un graphe d'ordre 5, il faut examiner $I + M + M^2 + M^3 + M^4$, et c'est bien le terme M^4 qui relie A à E .
- b) Les serveurs B, C et D sont tous critiques : chacun est un **point d'articulation**.
- Retirer C , par exemple, sépare $\{A; B\}$ de $\{D; E\}$ $\checkmark$.
 - A et E ne le sont pas : ce sont des extrémités, leur panne ne coupe pas les autres.

La correction. Ajouter l'arête EA transforme la chaîne en **cycle**.

- Chaque serveur a alors deux voisins, et la suppression de l'un quelconque laisse une chaîne connexe $\checkmark$.
- **Une seule arête suffit** donc, ce qui est le minimum : c'est exactement pourquoi les réseaux critiques sont câblés en anneau et non en ligne.

Réponse. a) oui · b) B, C et D sont critiques ; une seule arête (EA) suffit à fermer le réseau en anneau.

Correction de l'exercice 8 - Deux modes de transport

- a) Ici $a = 0,18$ et $b = 0,12$, l'état 1 étant « voiture ».
- Distribution invariante : $\frac{b}{a+b} = \frac{0,12}{0,30} = 0,4$.
 - **À long terme, 40 % des habitants prennent la voiture** et 60 % les transports en commun.
- b) La relation est $v_{n+1} = (1 - 0,18 - 0,12)v_n + 0,12 = 0,7v_n + 0,12$, avec $v_0 = 0,7$.
- Point fixe : $\ell = 0,7\ell + 0,12$ donne $0,3\ell = 0,12$, soit $\ell = 0,4 \checkmark$.
 - $v_n - 0,4$ est géométrique de raison 0,7, avec $v_0 - 0,4 = 0,3$:

$$v_n = 0,4 + 0,3 \times 0,7^n$$
 - **Vérification** : $v_1 = 0,4 + 0,21 = 0,61$, et le direct donne $0,7 \times 0,7 + 0,12 = 0,61 \checkmark$.
 - **Lecture pour la collectivité** : la bascule est lente (raison 0,7). Il faut environ $n = 8$ jours pour descendre sous 45 %, puisque $0,3 \times 0,7^8 \approx 0,017$.

Réponse. a) 40 % voiture, 60 % transports · b) $v_n = 0,4 + 0,3 \times 0,7^n$.

Correction de l'exercice 9 - Chemins dans un réseau orienté

- a) Un 1 en ligne i , colonne j pour chaque flux $i \rightarrow j$:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

— La matrice n'est pas symétrique : les flux sont orientés ✓.

b) On suit les flèches depuis le poste 1 :

— $1 \rightarrow 2 \rightarrow 4 \rightarrow 1$: c'est un **circuit de longueur 3**.

— $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$: c'est un **circuit de longueur 4**.

— Il y a donc deux circuits, de longueurs 3 et 4.

— **Vérification matricielle** : le coefficient $(1; 1)$ de A^3 vaut 1 (le circuit court) et celui de A^4 vaut 1 (le circuit long).

— **Lecture industrielle** : un circuit dans un graphe de production signale un **recyclage** — la matière repasse par le poste 1. Si le graphe représentait des dépendances de tâches, ce circuit rendrait l'ordonnancement impossible.

Réponse. a) A ci-dessus · b) oui : $1 \rightarrow 2 \rightarrow 4 \rightarrow 1$ (longueur 3) et $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$ (longueur 4).

Correction de l'exercice 10 - Un marcheur sur une étoile

a) Degrés : $\deg(C) = 4$ et $\deg(A_i) = 1$ pour chacun des quatre autres.

— Somme des degrés : $4 + 4 \times 1 = 8$, donc le graphe a 4 arêtes.

— Distribution invariante : $p_C = \frac{4}{8} = 0,5$ et $p_{A_i} = \frac{1}{8} = 0,125$.

$$P = (0,5; 0,125; 0,125; 0,125; 0,125)$$

— **Contrôle** : la somme vaut $0,5 + 4 \times 0,125 = 1$ ✓.

— **Interprétation** : le marcheur passe **la moitié** de son temps au centre — ce qui est évident, puisqu'il y revient à chaque pas depuis une branche.

b) **Non**, la marche ne converge pas.

— L'étoile est **biparte** : les groupes sont $\{C\}$ et $\{A_1, \dots, A_4\}$, et chaque arête relie l'un à l'autre.

— Le marcheur est donc au centre aux étapes paires et sur une branche aux étapes impaires (s'il part du centre) : la distribution alterne indéfiniment.

— **Ce qui reste vrai** : la distribution invariante décrit la **proportion de temps** passée à chaque sommet, et elle vaut bien $\frac{1}{2}$ pour le centre ✓.

Réponse. a) $p_C = 0,5$ et $p_{A_i} = 0,125$ · b) non : l'étoile est biparte, la marche oscille.

Correction de l'exercice 11 - Ajouter un lien change le classement

Ce qui change dans la matrice. La page 2 avait deux liens sortants (vers 1 et 3), chacun de poids $\frac{1}{2}$. Elle en a maintenant trois, chacun de poids $\frac{1}{3}$.

— La colonne 2 passe de $\left(\frac{1}{2}; 0; \frac{1}{2}; 0\right)$ à $\left(\frac{1}{3}; 0; \frac{1}{3}; \frac{1}{3}\right)$.

Les deux effets.

— **Effet positif pour la page 4** : elle reçoit désormais une part du « vote » de la page 2, qui pèse 0,3. Son score augmente donc d'environ $\frac{1}{3} \times 0,3 = 0,1$ au premier tour.

- **Effet négatif pour les pages 1 et 3** : le vote de la page 2 est maintenant partagé en trois au lieu de deux, elles perdent chacune un peu.

Conclusion. Le classement de la page 4 s'améliore nettement, au détriment des pages 1 et 3.

- **Le principe du PageRank** : le score d'une page est la somme des scores des pages qui pointent vers elle, chacun **divisé par leur nombre de liens sortants**. Recevoir un lien d'une page importante qui en émet peu vaut beaucoup plus que dix liens de pages sans audience.
- **La conséquence économique** : c'est toute l'industrie du référencement qui repose sur cette formule, et la raison pour laquelle l'achat de liens est traqué par les moteurs.

Réponse. La page 4 reçoit une part du score de la page 2 ; les pages 1 et 3 voient ce vote se diviser en trois au lieu de deux.

Correction de l'exercice 12 - Retour vers la matrice

On lit les trois informations de la formule.

- **La limite** est 0,25 : c'est $\frac{b}{a+b} = 0,25$, donc $b = 0,25(a+b)$, soit $3b = a$... vérifions : $b = 0,25a + 0,25b$ donne $0,75b = 0,25a$, donc $a = 3b$.
- **La raison** est 0,6 : $1 - a - b = 0,6$, donc $a + b = 0,4$.
- En combinant : $3b + b = 0,4$, donc $b = 0,1$ et $a = 0,3$.
- **La valeur initiale** : $p_0 = 0,25 + 0,55 = 0,8$.

$$T = \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix} = \begin{pmatrix} 0,7 & 0,1 \\ 0,3 & 0,9 \end{pmatrix}$$

- **Vérifications** : colonnes de somme 1 ✓ ; distribution invariante $\frac{0,1}{0,4} = 0,25$ ✓ ; raison $1 - 0,4 = 0,6$ ✓.
- **Contrôle sur un terme** : p_1 devrait valoir $0,7 \times 0,8 + 0,1 \times 0,2 = 0,56 + 0,02 = 0,58$, et la formule donne $0,25 + 0,55 \times 0,6 = 0,58$ ✓.

Réponse. $T = \begin{pmatrix} 0,7 & 0,1 \\ 0,3 & 0,9 \end{pmatrix}$ et $p_0 = 0,8$.

Correction de l'exercice 13 - Une file d'attente simplifiée

a) On prend l'état 1 = « libre ». Alors $a = 0,3$ (probabilité de devenir occupé) et $b = 0,5$ (probabilité de redevenir libre).

- Proportion de temps **libre** : $\frac{b}{a+b} = \frac{0,5}{0,8} = 0,625$.
- Proportion de temps **occupé** : $1 - 0,625 = 0,375$, soit 37,5 %.
- **Vérification par la formule symétrique** : $\frac{a}{a+b} = \frac{0,3}{0,8} = 0,375$ ✓.

b) Une journée de 8 heures compte $8 \times 60 = 480$ minutes.

- Temps occupé : $0,375 \times 480 = 180$ minutes, soit 3 heures.
- **Interprétation** : le guichet travaille trois heures sur huit. Un taux d'occupation de 37,5 % laisse une large marge, mais ce calcul **ne dit rien** de la longueur des files : un modèle à deux états ne compte pas les clients en attente.
- **Pour aller plus loin** : la durée moyenne d'un service est $\frac{1}{0,5} = 2$ minutes, et celle d'une période libre $\frac{1}{0,3} \approx 3,33$ minutes — le rapport $\frac{2}{2 + 3,33} = 0,375$ retrouve bien

le taux d'occupation ✓.

Réponse. a) 37,5 % du temps · b) 180 minutes, soit 3 heures.

Correction de l'exercice 14 - Comparer deux réseaux

Réseau A (chaîne). Degrés : 1, 2, 2, 2, 1, soit une somme de 8 et 4 arêtes.

— Distribution invariante : $\left(\frac{1}{8}; \frac{2}{8}; \frac{2}{8}; \frac{2}{8}; \frac{1}{8}\right)$.

— Les extrémités sont deux fois moins visitées que les sommets intérieurs.

Réseau B (cycle). Degrés : tous égaux à 2, somme 10, donc 5 arêtes.

— Distribution invariante : $\frac{2}{10} = \frac{1}{5}$ pour chaque sommet — **équirépartition**, conformément à la symétrie du cycle.

Robustesse.

— Dans la chaîne, la panne de S_2 , S_3 ou S_4 **coupe** le réseau en deux morceaux.

— Dans le cycle, la panne d'un sommet quelconque laisse une chaîne : le réseau **reste connexe**.

— **Le prix à payer** : une seule arête de plus ($S_5 - S_1$) pour passer de 4 à 5 liaisons.

Convergence. Aucun des deux ne garantit la convergence de la même façon :

— La chaîne est **biparte** : la marche oscille.

— Le cycle de longueur 5 contient un cycle **impair** : la marche y **converge** vers l'équirépartition.

— **Conclusion** : le cycle est meilleur sur les trois critères — équité, robustesse et convergence.

Réponse. Chaîne : $\left(\frac{1}{8}; \frac{2}{8}; \frac{2}{8}; \frac{2}{8}; \frac{1}{8}\right)$, fragile et biparte · Cycle : équirépartition $\frac{1}{5}$, robuste, convergent.

Correction de l'exercice 15 - Bilan : un réseau de bout en bout

a) Avec l'ordre A, B, C, D :

$$M = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

— Degrés : $\deg(A) = 2$, $\deg(B) = 3$, $\deg(C) = 2$, $\deg(D) = 1$.

— Somme 8, donc 4 arêtes ✓ — c'est bien le nombre annoncé.

b) On calcule le coefficient $(1; 4)$ de M^2 : ligne 1 de M contre colonne 4 :

— $0 \times 0 + 1 \times 1 + 1 \times 0 + 0 \times 0 = 1$.

— Il y a **un seul** chemin de longueur 2 de A à D : $A - B - D$.

— C'est cohérent : D n'a que B pour voisin, et A est bien relié à B ✓.

c) $p_i = \frac{\deg(i)}{8}$:

$$P = \left(\frac{2}{8}; \frac{3}{8}; \frac{2}{8}; \frac{1}{8}\right) = (0,25; 0,375; 0,25; 0,125)$$

— **Vérification** de la ligne A : le marcheur arrive en A depuis B (probabilité $\frac{1}{3}$) ou depuis C (probabilité $\frac{1}{2}$), soit $\frac{1}{3} \times 0,375 + \frac{1}{2} \times 0,25 = 0,125 + 0,125 = 0,25$ ✓.

- **Convergence** : le graphe est connexe et contient le triangle ABC , donc un cycle de longueur **impair** : il n'est pas biparti, la marche **converge** vers P .
- **Interprétation** : le sommet B , seul point d'articulation du réseau, est aussi le plus visité (37,5 %) ; le sommet D , en bout de branche, l'est le moins (12,5 %).

Réponse. a) degrés 2, 3, 2, 1 · b) un seul : $A - B - D$ · c) (0,25 ; 0,375 ; 0,25 ; 0,125), convergence assurée (triangle ABC).