



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 1

2004

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Partie I : Étude d'une suite de vecteurs

I.1.a. Matrice de rang un et diagonalisation

Le coefficient de ligne i et colonne j du produit $C^t C$ est $c_i c_j$. Ainsi

$$C^t C = \begin{pmatrix} c_1^2 & c_1 c_2 & \cdots & c_1 c_n \\ c_2 c_1 & c_2^2 & \cdots & c_2 c_n \\ \vdots & \vdots & \ddots & \vdots \\ c_n c_1 & c_n c_2 & \cdots & c_n^2 \end{pmatrix}.$$

Cette matrice est réelle et symétrique, car $c_i c_j = c_j c_i$. Le théorème spectral réel s'applique : elle est diagonalisable dans une base orthonormale de $\mathbb{R}^n$.

I.1.b. Carré de la matrice

L'associativité du produit matriciel donne

$$(C^t C)^2 = C(t C C)^t C.$$

Or

$$t C C = \sum_i c_i^2 = \|C\|^2$$

est un scalaire. Donc

$$(C^t C)^2 = \|C\|^2 C^t C.$$

I.1.c. Valeurs propres possibles

Posons $M = C^t C$. Si $X \neq 0$ est un vecteur propre associé à μ , alors $MX = \mu X$ et $M^2 X = \mu^2 X$. L'identité de I.1.b donne également $M^2 X = \|C\|^2 \mu X$.

Ainsi $(\mu^2 - \|C\|^2 \mu)X = 0$. Comme $X \neq 0$, $\mu(\mu - \|C\|^2) = 0$, donc $\mu = 0$ ou $\mu = \|C\|^2$. On a seulement utilisé l'action d'un polynôme annulateur sur un vecteur propre.

I.1.d. Noyau et espaces propres

Question adaptée au programme actuel : le noyau est demandé explicitement pour couvrir aussi $n = 1$, cas où 0 n'est pas une valeur propre.

Pour tout $X \in \mathbb{R}^n$,

$$MX = C({}^tCX) = \langle C, X \rangle C.$$

Puisque $C \neq 0$, $MX = 0$ si et seulement si $\langle C, X \rangle = 0$. Donc

$$\text{Ker } M = \text{Vect}(C)^\perp.$$

Ce sous-espace a dimension $n - 1$. Si $n \geq 2$, il contient un vecteur non nul : 0 est une valeur propre et $E_0(M) = \text{Vect}(C)^\perp$. Si $n = 1$, le noyau est $\{0\}$ et 0 n'est pas une valeur propre.

En appliquant la formule à C ,

$$MC = \langle C, C \rangle C = \|C\|^2 C.$$

La valeur $\|C\|^2 > 0$ est donc toujours propre. Si $MX = \|C\|^2 X$, la formule de MX montre que X appartient à $\text{Vect}(C)$. Réciproquement, tous les multiples de C vérifient cette égalité. Ainsi

$$E_{\|C\|^2}(M) = \text{Vect}(C).$$

I.1.e. Interprétation géométrique

La projection orthogonale Π_C sur la droite $\text{Vect}(C)$ est donnée par

$$\Pi_C(X) = \frac{\langle C, X \rangle}{\|C\|^2} C,$$

le dénominateur étant strictement positif. La formule de I.1.d donne donc

$$MX = \|C\|^2 \Pi_C(X).$$

L'endomorphisme associé à M est une projection orthogonale sur cette droite suivie d'une multiplication par $\|C\|^2$. Il a pour image $\text{Vect}(C)$ et pour noyau son orthogonal.

Lorsque C est unitaire, $\|C\|^2 = 1$ et M représente exactement la projection orthogonale Π_C . On peut aussi le vérifier par $M^2 = M$ et par la symétrie de M , qui caractérisent un

projecteur orthogonal.

I.2.a. Identités de produits

Le produit tXY est le scalaire

$$\sum_i x_i y_i,$$

donc ${}^tXY = {}^tYX$.

Par définition du produit scalaire, ${}^tXAY = \langle X, AY \rangle$. Comme ${}^tA = A$,

$$\langle AX, Y \rangle = {}^t(AX)Y = {}^tX{}^tAY = {}^tXAY.$$

Pour les dernières identités, l'associativité donne

$${}^tY(X{}^tY)X = ({}^tYX)({}^tYX) = ({}^tXY)^2,$$

$${}^tY(X{}^tX)Y = ({}^tYX)({}^tXY) = ({}^tXY)^2.$$

Toutes les tailles sont compatibles : les matrices entre parenthèses sont d'ordre n et les produits finaux sont des scalaires.

I.2.b. Base orthonormale de vecteurs propres

La matrice A est réelle et symétrique. Le théorème spectral réel garantit une base orthonormale $(U_1, \dots, U_n)$ de $\mathbb{R}^n$ constituée de vecteurs propres de A .

Pour chaque i , il existe donc un réel λ_i tel que $AU_i = \lambda_i U_i$. Les valeurs propres peuvent être répétées dans cette liste, selon la dimension de leurs espaces propres.

I.2.c. Coordonnées, normes et forme quadratique

Posons $\xi_i = \langle U_i, X \rangle$. Dans une base orthonormale, les coordonnées d'un vecteur sont ses produits scalaires avec les vecteurs de base. Ainsi

$$X = \sum_{i=1}^n \xi_i U_i, \quad \|X\|^2 = \sum_{i=1}^n \xi_i^2.$$

Par linéarité et puisque $AU_i = \lambda_i U_i$,

$$AX = \sum_{i=1}^n \lambda_i \xi_i U_i.$$

La symétrie établie en I.2.a donne aussi $\langle U_i, AX \rangle = \langle AU_i, X \rangle = \lambda_i \xi_i$. Donc

$$AX = \sum_{i=1}^n \langle U_i, AX \rangle U_i, \quad \|AX\|^2 = \sum_{i=1}^n \langle U_i, AX \rangle^2 = \sum_{i=1}^n \lambda_i^2 \xi_i^2.$$

Les normes sont les racines carrées positives de ces sommes. Enfin, l'orthonormalité donne

$$\langle X, AX \rangle = \sum_{i=1}^n \sum_{j=1}^n \xi_i \lambda_j \xi_j \langle U_i, U_j \rangle = \sum_{i=1}^n \lambda_i \xi_i^2.$$

C'est l'égalité demandée.

I.2.d. Décompositions matricielles

Pour tout $X \in \mathbb{R}^n$, $U_i^t(U_i)X = \langle U_i, X \rangle U_i$. Par les décompositions de I.2.c,

$$\left(\sum_{i=1}^n U_i^t(U_i) \right) X = X,$$

$$\left(\sum_{i=1}^n \lambda_i U_i^t(U_i) \right) X = AX.$$

Deux matrices qui agissent de la même manière sur tous les vecteurs, en particulier sur

la base canonique, sont égales. Ainsi

$$I = \sum_{i=1}^n U_i {}^t(U_i), \quad A = \sum_{i=1}^n \lambda_i U_i {}^t(U_i).$$

Comme chaque U_i est unitaire, I.1.e montre que $U_i {}^t(U_i)$ représente la projection orthogonale sur $\text{Vect}(U_i)$.

I.2.e. Encadrement de la forme quadratique

Notons $m = \min_i \lambda_i$ et $M = \max_i \lambda_i$. Pour chaque i , le nombre ξ_i^2 est positif ou nul et

$$m\xi_i^2 \leq \lambda_i \xi_i^2 \leq M\xi_i^2.$$

En sommant puis en utilisant I.2.c,

$$m\|X\|^2 \leq \langle X, AX \rangle \leq M\|X\|^2.$$

Ces inégalités valent aussi pour $X = 0$, où tous les termes sont nuls.

I.2.f. Application à la matrice tridiagonale

Méthode 1 — Encadrement de la forme quadratique

Pour $X = {}^t(x_1, \dots, x_n)$, le produit matriciel donne

$$\langle X, AX \rangle = 4 \sum_{i=1}^n x_i^2 - 2 \sum_{i=1}^{n-1} x_i x_{i+1}.$$

La somme de produits est vide si $n = 1$. Pour tous réels a, b , $2|ab| \leq a^2 + b^2$, puisque $(|a| - |b|)^2 \geq 0$. Il en résulte

$$\left| 2 \sum_{i=1}^{n-1} x_i x_{i+1} \right| \leq \sum_{i=1}^{n-1} (x_i^2 + x_{i+1}^2) \leq 2\|X\|^2.$$

Ainsi $2\|X\|^2 \leq \langle X, AX \rangle \leq 6\|X\|^2$. Pour un vecteur propre non nul X associé à λ , $\langle X, AX \rangle = \lambda\|X\|^2$. La division par $\|X\|^2 > 0$ donne $2 \leq \lambda \leq 6$.

Méthode 2 — Une composante de valeur absolue maximale

Soit $X \neq 0$ un vecteur propre associé à λ . Choisissons un indice k tel que $|x_k| = \max_i |x_i| > 0$, et posons $x_0 = x_{n+1} = 0$ pour traiter les extrémités.

La ligne k de $AX = \lambda X$ donne $(4 - \lambda)x_k = x_{k-1} + x_{k+1}$. Donc

$$|4 - \lambda||x_k| \leq |x_{k-1}| + |x_{k+1}| \leq 2|x_k|.$$

En divisant par $|x_k| > 0$, $|4 - \lambda| \leq 2$, d'où le même encadrement $2 \leq \lambda \leq 6$. La première méthode contrôle tous les vecteurs ; la seconde exploite directement l'équation propre, sans théorème supplémentaire sur la localisation des valeurs propres.

On peut même préciser que les bornes sont strictes. En effet,

$$\langle X, (A - 2I)X \rangle = x_1^2 + x_n^2 + \sum_{i=1}^{n-1} (x_i - x_{i+1})^2,$$

$$\langle X, (6I - A)X \rangle = x_1^2 + x_n^2 + \sum_{i=1}^{n-1} (x_i + x_{i+1})^2.$$

Pour $X \neq 0$, chacune de ces sommes est strictement positive : si elle était nulle, les extrémités seraient nulles et les relations successives imposeraient toutes les composantes nulles. Ainsi les valeurs propres appartiennent en fait à $]2, 6[$.

I.3.a. Norme de AX

L'égalité

$$\|AX\|^2 = \sum_i \lambda_i^2 \langle U_i, X \rangle^2$$

a été établie en I.2.c. Comme $|\lambda_i| \leq \rho(A)$,

$$\|AX\|^2 \leq \rho(A)^2 \sum_i \langle U_i, X \rangle^2 = \rho(A)^2 \|X\|^2.$$

Les deux côtés étant positifs ou nuls, la prise de racines carrées donne $\|AX\| \leq \rho(A)\|X\|$.

La liste des valeurs propres est finie et non vide ; il existe j tel que $|\lambda_j| = \rho(A)$. Le vecteur unitaire U_j réalise l'égalité, puisque $\|AU_j\| = |\lambda_j| = \rho(A)\|U_j\|$. Cela reste vrai si $\rho(A) = 0$.

I.3.b. Convergence des puissances

Méthode 1 — Itération de l'inégalité de norme

Si $\rho(A) < 1$, I.3.a appliquée successivement donne $\|A^p X\| \leq \rho(A)^p \|X\|$ pour $p \geq 1$. Si $\rho(A) = 0$, $A = 0$ par la base propre et tous ces vecteurs sont nuls ; sinon la suite géométrique $\rho(A)^p$ tend vers zéro. Donc $A^p X \rightarrow 0$ pour tout X .

Réciproquement, supposons $A^p X \rightarrow 0$ pour tout X . Pour chaque U_i , $A^p U_i = \lambda_i^p U_i$, donc $\|A^p U_i\| = |\lambda_i|^p \rightarrow 0$. Une suite géométrique de raison positive ou nulle tend vers zéro seulement si sa raison est strictement inférieure à 1. Ainsi tous les $|\lambda_i| < 1$, et leur maximum fini vérifie $\rho(A) < 1$.

Méthode 2 — Coordonnées dans la base propre

Si

$$X = \sum_i \xi_i U_i,$$

alors

$$A^p X = \sum_i \lambda_i^p \xi_i U_i$$

pour $p \geq 1$ et

$$\|A^p X\|^2 = \sum_i |\lambda_i|^{2p} \xi_i^2.$$

Si tous les $|\lambda_i| < 1$, chaque terme tend vers zéro ; la somme étant finie, la norme tend vers zéro. Si au moins un $|\lambda_j| \geq 1$, le choix $X = U_j$ donne $\|A^p X\| = |\lambda_j|^p \geq 1$, ce qui interdit la convergence vers zéro pour tout X .

On retrouve l'équivalence (i)–(ii). La première méthode donne une borne globale ; la seconde identifie précisément le comportement de chaque composante propre.

Partie II : Un problème de minimisation

II.1.a. Degré et coefficient dominant

On a $T_0 = 1$, de degré 0 et de coefficient dominant 1, puis $T_1(t) = t$, de degré 1 et de coefficient dominant $1 = 2^0$.

Supposons, pour un entier $p \geq 1$, que T_p ait degré p et coefficient dominant 2^{p-1} , tandis que T_{p-1} a degré $p-1$. Dans

$$T_{p+1}(t) = 2tT_p(t) - T_{p-1}(t),$$

le premier terme a degré $p + 1$ et coefficient dominant 2^p . Le second terme a un degré strictement inférieur et ne peut supprimer ce coefficient. Donc T_{p+1} a degré $p + 1$ et coefficient dominant 2^p .

Par récurrence, $\deg T_p = p$ pour tout $p \geq 0$, et le coefficient de t^p vaut 1 si $p = 0$ et 2^{p-1} si $p \geq 1$.

II.1.b. Identité trigonométrique

On raisonne par récurrence à deux rangs. Pour $p = 0$, $T_0(\cos \theta) = 1 = \cos 0$; pour $p = 1$, $T_1(\cos \theta) = \cos \theta$.

La formule d'addition donnée, appliquée à $p\theta + \theta$ et à $p\theta - \theta$, puis additionnée, donne

$$\cos((p+1)\theta) + \cos((p-1)\theta) = 2 \cos \theta \cos(p\theta).$$

Si l'identité est vraie aux rangs p et $p-1$, alors

$$T_{p+1}(\cos \theta) = 2 \cos \theta \cos(p\theta) - \cos((p-1)\theta) = \cos((p+1)\theta).$$

Ainsi $T_p(\cos \theta) = \cos(p\theta)$ pour tout réel θ et tout entier $p \geq 0$.

II.1.c. Maximum et racines

Question adaptée au programme actuel : les formules de racines sont réservées à $p \geq 1$, et le cas $p = 0$ est précisé séparément.

La fonction cosinus réalise une bijection de $[0, \pi]$ sur $[-1, 1]$. Pour tout t de cet intervalle, il existe donc $\theta \in [0, \pi]$ tel que $t = \cos \theta$. Par II.1.b, $|T_p(t)| = |\cos(p\theta)| \leq 1$. Comme $T_p(1) = 1$, le maximum, et donc le supremum, vaut 1.

Pour $p \geq 1$, les p angles

$$\theta_k = \frac{(2k+1)\pi}{2p}, \quad 0 \leq k \leq p-1,$$

appartiennent à $]0, \pi[$ et vérifient $\cos(p\theta_k) = 0$. Les nombres

$$t_k = \cos \left(\frac{(2k+1)\pi}{2p} \right)$$

sont donc des racines de T_p dans $] -1, 1[$. Ils sont distincts puisque le cosinus est strictement décroissant sur $[0, \pi]$. Un polynôme de degré p ne peut avoir plus de p racines distinctes ; ce sont donc exactement toutes ses racines, et elles sont simples.

Pour $p = 0$, $T_0 = 1$ n'a aucune racine et son maximum absolu vaut encore 1.

Une conséquence utile est que, pour $p \geq 1$,

$$T_p(t) = 2^{p-1} \prod_{k=0}^{p-1} (t - t_k).$$

Ainsi $T_p(a) \neq 0$ pour $|a| > 1$. Son signe est positif si $a > 1$ et vaut $(-1)^p$ si $a < -1$.

II.2.a. Alternance et contradiction

Dans cette question, $p \geq 1$: cette précision évite toute division par zéro dans les nœuds $\cos(j\pi/p)$.

Posons $s_j = \cos(j\pi/p)$, pour $0 \leq j \leq p$. On a $1 = s_0 > s_1 > \dots > s_p = -1$ et, par II.1.b,

$$S_p(s_j) = \frac{T_p(s_j)}{T_p(a)} = \frac{(-1)^j}{T_p(a)}.$$

Notons $M_p = 1/|T_p(a)| > 0$. Si le polynôme P satisfait l'inégalité stricte du sujet, $|P(s_j)| < M_p = |S_p(s_j)|$. La différence $D = S_p - P$ possède donc au nœud s_j le même signe strict que $S_p(s_j)$:

$$\text{sgn } D(s_j) = (-1)^j \text{sgn } T_p(a).$$

Ces signes alternent entre chaque paire de nœuds voisins. Le polynôme D est continu ; le théorème des valeurs intermédiaires donne une racine dans chaque intervalle $]s_{j+1}, s_j[$, pour $0 \leq j < p$. Ces p racines sont distinctes.

De plus, $D(a) = S_p(a) - P(a) = 1 - 1 = 0$. Comme $|a| > 1$, cette racine est extérieure à $[-1, 1]$ et distincte des précédentes. Ainsi D possède au moins $p + 1$ racines distinctes.

Or $\deg D \leq p$ et $D \neq 0$, puisque ses valeurs aux nœuds sont non nulles. Cela contredit le nombre maximal de racines d'un polynôme non nul. Aucun tel P n'existe.

II.2.b. Valeur du minimum

Le polynôme S_p appartient à $\mathbb{R}_p[X]$, satisfait $S_p(a) = 1$ et, par II.1.c,

$$\sup_{-1 \leq t \leq 1} |S_p(t)| = \frac{1}{|T_p(a)|} = M_p.$$

La question II.2.a montre qu'aucun polynôme admissible ne peut avoir un supremum strictement plus petit. Ainsi S_p atteint le minimum et sa valeur est M_p .

Les supremums utilisés sont des maxima : les polynômes sont continus sur le segment $[-1, 1]$.

II.2.c. Unicité du polynôme minimisant

Soit P un minimiseur. Alors $P(a) = 1$ et $\sup_{[-1, 1]} |P| = M_p$. Posons $R = (P + S_p)/2$. On a $R(a) = 1$ et, pour tout $t \in [-1, 1]$,

$$|R(t)| \leq \frac{|P(t)| + |S_p(t)|}{2} \leq M_p.$$

Par la minimalité prouvée en II.2.b, son supremum vaut exactement M_p ; R est également minimiseur.

Il faut encore établir que R atteint cette borne à chaque nœud s_j , ce qui ne découle pas du seul fait que son supremum vaut M_p . On le démontre par une identité polynomiale élémentaire.

Pour $0 \leq j \leq p$, définissons

$$L_j(t) = \prod_{\substack{0 \leq k \leq p \\ k \neq j}} \frac{t - s_k}{s_j - s_k}.$$

Les nœuds sont distincts, donc les dénominateurs sont non nuls. Les polynômes L_j ont degré p et vérifient $L_j(s_k) = 1$ si $k = j$, et 0 sinon. Pour tout polynôme D de degré au plus p , le polynôme

$$D(t) - \sum_{j=0}^p D(s_j) L_j(t)$$

s'annule aux $p + 1$ nœuds et a degré au plus p ; il est donc nul. Cette démonstration établit l'identité d'interpolation nécessaire.

Appliquons-la à $D = S_p - R$. Comme $D(a) = 0$, on a

$$0 = \sum_{j=0}^p D(s_j) L_j(a).$$

À chaque nœud, $|R(s_j)| \leq M_p$ et $S_p(s_j)$ vaut soit M_p , soit $-M_p$. Donc $D(s_j)$ est de même signe que $S_p(s_j)$, ou nul.

Déterminons maintenant le signe de $L_j(a)$. Les nœuds étant décroissants, le dénominateur

$$\prod_{k \neq j} (s_j - s_k)$$

possède exactement j facteurs négatifs; son signe est $(-1)^j$. Si $a > 1$, tous les facteurs du numérateur sont positifs, donc $\text{sgn } L_j(a) = (-1)^j$. Si $a < -1$, ses p facteurs sont négatifs, donc $\text{sgn } L_j(a) = (-1)^{p+j}$.

Par le signe de $T_p(a)$ déterminé en II.1.c, c'est dans les deux cas exactement le signe de $S_p(s_j) = (-1)^j / T_p(a)$. Chaque produit $D(s_j) L_j(a)$ est donc positif ou nul. Leur somme vaut zéro, et tous les $L_j(a)$ sont non nuls; il en résulte $D(s_j) = 0$ pour chaque j .

Ainsi $R(s_j) = S_p(s_j)$ et

$$\frac{1}{2} |P(s_j) + S_p(s_j)| = |R(s_j)| = M_p = \frac{1}{|T_p(a)|}.$$

C'est l'égalité demandée. Enfin, $R - S_p$ a degré au plus p et s'annule aux $p + 1$ nœuds distincts, donc $R = S_p$. Comme $R = (P + S_p)/2$, on obtient $P = S_p$. Le minimiseur est unique.

II.3. Transport sur $[\alpha, \beta]$

La transformation affine

$$z = \frac{2t - \alpha - \beta}{\beta - \alpha}$$

est une bijection de $[\alpha, \beta]$ sur $[-1, 1]$, de réciproque $t = (\alpha + \beta)/2 + (\beta - \alpha)z/2$. Comme $\beta - \alpha > 0$, elle est bien définie.

La valeur correspondant à $t = 0$ est

$$a = \frac{-\alpha - \beta}{\beta - \alpha} = \frac{\alpha + \beta}{\alpha - \beta} < -1,$$

puisque $\alpha > 0$. À tout polynôme Q de degré au plus p , associons

$$R(z) = Q\left(\frac{\alpha + \beta}{2} + \frac{\beta - \alpha}{2}z\right).$$

Cette correspondance est bijective sur les polynômes de degré au plus p . Elle préserve les supremums sur les segments correspondants et transforme $Q(0) = 1$ en $R(a) = 1$.

Pour $p \geq 1$, les questions II.2.b–II.2.c donnent l'unique minimiseur $R = T_p/T_p(a)$. En revenant à t , on obtient

$$Q_*(t) = \frac{T_p\left(\frac{2t - \alpha - \beta}{\beta - \alpha}\right)}{T_p\left(\frac{\alpha + \beta}{\alpha - \beta}\right)}.$$

Son dénominateur est non nul car $|a| > 1$; sa valeur minimale est $1/|T_p(a)|$.

Question adaptée au programme actuel : le cas $p = 0$ est explicité. Un polynôme constant vérifiant $Q(0) = 1$ est nécessairement $Q = 1$. C'est donc l'unique solution, de valeur minimale 1. La formule affichée reste valable pour $p = 0$, puisque $T_0 = 1$.

Partie III : Résolution itérative du système

Préambule — Existence et unicité de X^*

Les valeurs propres de A sont toutes strictement positives, donc non nulles. Dans la base orthonormale propre de I.2.b, la matrice de A est diagonale avec ces valeurs sur la diagonale ; elle est donc inversible.

Ainsi, pour tout $B \in \mathbb{R}^n$, le système $AX = B$ possède une unique solution $X^* = A^{-1}B$. Plus explicitement, en écrivant

$$B = \sum_i b_i U_i,$$

$$X^* = \sum_{i=1}^n \frac{b_i}{\lambda_i} U_i,$$

les divisions étant licites puisque $\lambda_i > 0$.

III.1.a. Expression de l'erreur pour un pas constant

Posons $E_p = X_p - X^*$. Comme $AX^* = B$ et $\alpha_p = \alpha$,

$$E_{p+1} = X_p - X^* + \alpha(B - AX_p) = E_p - \alpha AE_p = (I - \alpha A)E_p.$$

La récurrence commence en $E_0 = X_0 - X^*$. Une récurrence sur p donne donc

$$\forall p \in \mathbb{N}, \quad X_p - X^* = (I - \alpha A)^p (X_0 - X^*).$$

Pour $p = 0$, la puissance de la matrice est l'identité.

III.1.b. Rayon spectral et tracé

Dans la base propre de A , on a

$$(I - \alpha A)U_i = (1 - \alpha \lambda_i)U_i.$$

Comme ces vecteurs forment une base, les valeurs propres, avec leurs répétitions, sont $\mu_i = 1 - \alpha\lambda_i$. La matrice $I - \alpha A$ est symétrique réelle et

$$f(\alpha) = \rho(I - \alpha A) = \max_i |1 - \alpha\lambda_i|.$$

Notons désormais $m = \lambda_1 > 0$ et $M = \lambda_n \geq m$. Pour tout $\lambda \in [m, M]$ lorsque $m < M$, on peut écrire $\lambda = (1 - s)m + sM$ avec $0 \leq s \leq 1$. L'inégalité triangulaire donne

$$|1 - \alpha\lambda| \leq (1 - s)|1 - \alpha m| + s|1 - \alpha M| \leq \max(|1 - \alpha m|, |1 - \alpha M|).$$

Les deux extrémités sont elles-mêmes des valeurs propres, donc

$$f(\alpha) = \max(|1 - \alpha m|, |1 - \alpha M|).$$

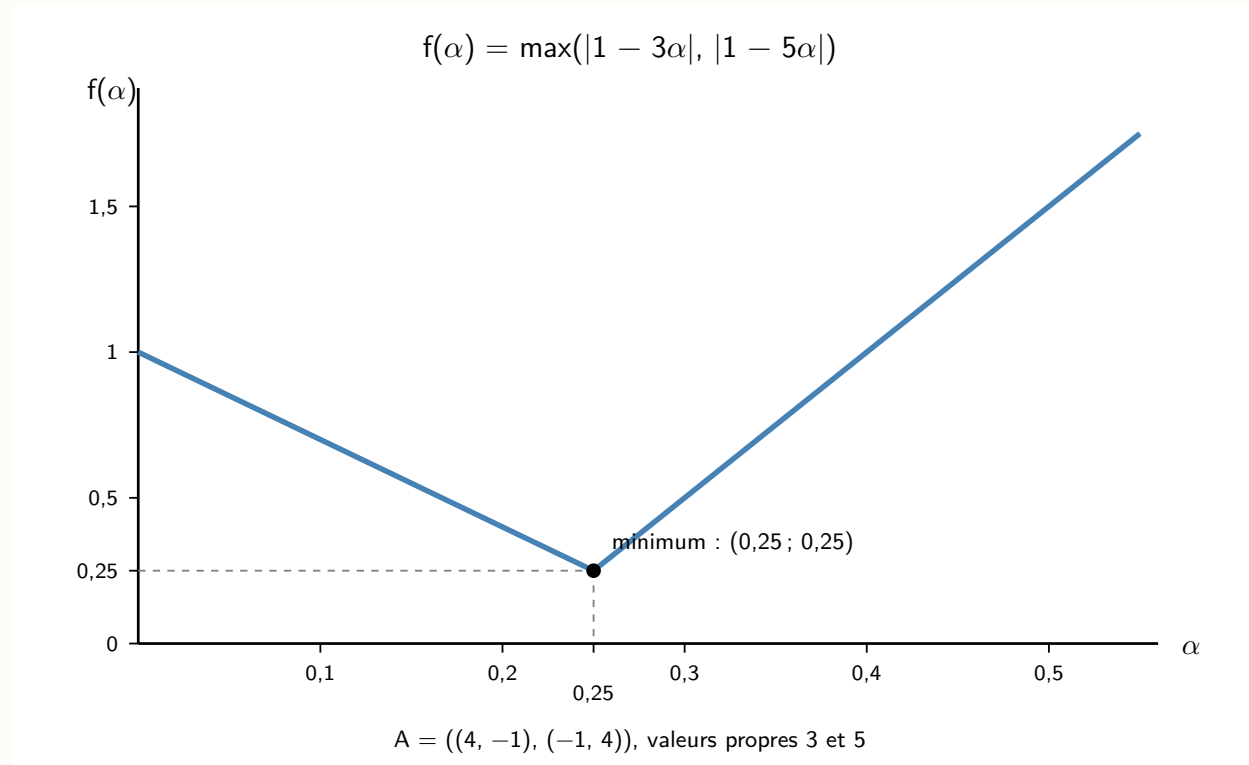
La formule vaut aussi si $m = M$. Posons $\alpha_* = 2/(m + M)$ et $r = (M - m)/(M + m)$.

Pour $0 \leq \alpha \leq \alpha_*$, $1 - \alpha m \geq 0$ et domine $1 - \alpha M$ ainsi que $\alpha M - 1$. Pour $\alpha \geq \alpha_*$, $\alpha M - 1 \geq 0$ et domine les deux signes de $1 - \alpha m$. Ainsi

$$f(\alpha) = \begin{cases} 1 - \alpha m & 0 \leq \alpha \leq \alpha_*, \\ \alpha M - 1 & \alpha \geq \alpha_*. \end{cases}$$

Le tracé est donc formé de deux demi-segments de droites : il part de $(0, 1)$, décroît avec pente $-m$ jusqu'au point (α_*, r) , puis croît avec pente M . Il est entièrement déterminé par cette formule, pour toutes les valeurs de m et M .

Pour rendre le tracé concret, la figure ci-dessous prend la matrice déjà définie en I.2.f dans le cas $n = 2$. Ses vecteurs ${}^t(1, 1)$ et ${}^t(1, -1)$ ont respectivement pour valeurs propres 3 et 5. Le tracé représente donc exactement $f(\alpha) = \max(|1 - 3\alpha|, |1 - 5\alpha|)$, de minimum $1/4$ en $\alpha = 1/4$.



III.1.c. Convergence et pas constant optimal

Question adaptée au programme actuel : l'équivalence est formulée pour tout vecteur initial. Pour un seul vecteur initial fixé, elle peut être fautive ; par exemple $X_0 = X^*$ donne une suite constante égale à X^* quel que soit α .

Par III.1.a et I.3.b appliquée à la matrice symétrique $I - \alpha A$, la convergence vers X^* pour tout X_0 équivaut à $\rho(I - \alpha A) < 1$. Comme les λ_i sont strictement positives,

$$|1 - \alpha\lambda_i| < 1 \iff 0 < \alpha\lambda_i < 2.$$

Sous l'hypothèse $\alpha > 0$, ces conditions pour tous les indices équivalent à $0 < \alpha < 2/M = 2/\lambda_n$.

Le sens de l'optimalité est la minimisation du facteur de contraction en norme, uniformément sur l'erreur initiale. Par I.3.a, $f(\alpha)$ est le meilleur coefficient dans $\|(I - \alpha A)E\| \leq f(\alpha)\|E\|$, puisqu'un vecteur propre le réalise. Le tracé de III.1.b montre que son unique minimum est atteint en

$$\alpha_* = \frac{2}{m + M}, \quad f(\alpha_*) = r = \frac{M - m}{M + m}.$$

Si $m < M$, on a $0 < r < 1$, et l'itération de la borne donne

$$\forall p \geq 1, \quad \|X_p - X^*\| \leq r^p \|X_0 - X^*\|.$$

En fait, pour le pas optimal, la matrice de l'erreur a les valeurs propres r et $-r$ sur les espaces correspondant aux extrémités ; la borne est atteinte pour une erreur initiale dans l'un de ces espaces. Elle est donc optimale comme borne uniforme.

Pour $p = 0$, l'inégalité se réduit à l'égalité $\|X_0 - X^*\| = \|X_0 - X^*\|$.

Si $m = M$, toutes les valeurs propres de A valent m . Comme A est diagonalisable, $A = mI$. Le pas optimal est $\alpha_* = 1/m$, d'où $I - \alpha_* A = 0$ et $X_1 = X^*$. Toutes les erreurs aux rangs $p \geq 1$ sont nulles. Cela traite ce cas sans utiliser une expression ambiguë de type 0^0 .

Pour un vecteur initial particulier, on peut préciser la condition exacte : si

$$X_0 - X^* = \sum_i e_i U_i,$$

la convergence a lieu si et seulement si $|1 - \alpha \lambda_i| < 1$ pour chaque indice dont le coefficient e_i est non nul. Les composantes absentes ne contraignent pas le pas.

III.2.a. Valeurs propres du polynôme de matrice

Les matrices $I - \alpha_j A$ sont des polynômes en A , donc commutent entre elles. Dans la base propre de A , chacune est diagonale et agit sur U_i par multiplication par $1 - \alpha_j \lambda_i$.

Ainsi

$$P_p(A)U_i = \left(\prod_{j=0}^{p-1} (1 - \alpha_j \lambda_i) \right) U_i = P_p(\lambda_i)U_i.$$

Les valeurs propres, avec leurs répétitions, sont donc $\nu_i = P_p(\lambda_i)$. La matrice $P_p(A)$ est symétrique : elle est une combinaison linéaire réelle des puissances de la matrice symétrique A , chacune de ces puissances étant symétrique.

On obtient

$$\rho(P_p(A)) = \max_i |P_p(\lambda_i)|.$$

Chaque λ_i appartient au segment $[m, M]$, donc

$$\rho(P_p(A)) \leq \sup_{m \leq t \leq M} |P_p(t)|.$$

Le supremum est fini et atteint, puisque P_p est continu sur ce segment, y compris si le segment se réduit à un point.

III.2.b. Expression et majoration de l'erreur

Pour un pas variable, la relation de l'erreur est

$$E_{j+1} = (I - \alpha_j A) E_j.$$

En la répétant pour les rangs 0 à $p - 1$, on obtient

$$E_p = (I - \alpha_{p-1} A) \cdots (I - \alpha_0 A) E_0.$$

Les facteurs commutent, donc ce produit est $P_p(A)$ tel qu'il est défini dans le sujet. Ainsi $X_p - X^* = P_p(A)(X_0 - X^*)$.

Appliquons I.3.a à la matrice symétrique $P_p(A)$, puis III.2.a :

$$\|X_p - X^*\| \leq \rho(P_p(A)) \|X_0 - X^*\| \leq \sup_{m \leq t \leq M} |P_p(t)| \|X_0 - X^*\|.$$

C'est la majoration demandée. Elle ne suppose pas que chacun des pas soit individuellement optimal.

III.2.c. Choix des pas, équivalent rectifié et comparaison

Question adaptée au programme actuel : on distingue $m < M$ du cas $m = M$, et on rectifie l'équivalent asymptotique erroné du PDF.

Choix optimal lorsque $m < M$

On fixe un entier $p \geq 1$. Le polynôme

$$P_p(t) = \prod_{j=0}^{p-1} (1 - \alpha_j t)$$

a degré p puisque tous les $\alpha_j > 0$, et vérifie $P_p(0) = 1$.

Par II.3, parmi tous les polynômes de degré au plus p ayant cette valeur en zéro, le minimiseur unique sur $[m, M]$ est

$$Q_*(t) = \frac{T_p\left(\frac{2t - m - M}{M - m}\right)}{T_p\left(\frac{m + M}{m - M}\right)}.$$

Il reste à vérifier qu'il peut réellement s'écrire sous la forme imposée avec des pas positifs. Ses racines, par II.1.c et la transformation affine, sont

$$\tau_j = \frac{m + M}{2} + \frac{M - m}{2} \cos\left(\frac{(2j + 1)\pi}{2p}\right), \quad 0 \leq j \leq p - 1.$$

Elles sont distinctes et appartiennent à $]m, M[$, donc sont toutes strictement positives. Choisissons

$$\alpha_j = \frac{1}{\tau_j} = \frac{2}{m + M + (M - m) \cos\left(\frac{(2j + 1)\pi}{2p}\right)}.$$

Les dénominateurs sont strictement positifs. Le produit

$$\prod_j (1 - t/\tau_j)$$

et Q_* ont le même degré, les mêmes racines et valent tous deux 1 en zéro. Ils sont donc égaux : leur factorisation par ces p racines ne laisse qu'un facteur constant, fixé par la valeur en zéro.

Le polynôme minimisant est ainsi réalisable par des pas positifs. Leur ordre peut être permuté, car les facteurs commutent ; le polynôme final reste le même. La valeur minimale et la borne sur l'erreur sont

$$\min \sup_{m \leq t \leq M} |P_p(t)| = \frac{1}{\left|T_p\left(\frac{m + M}{m - M}\right)\right|},$$

$$\|X_p - X^*\| \leq \frac{1}{\left|T_p\left(\frac{m+M}{m-M}\right)\right|} \|X_0 - X^*\|.$$

Expression exacte et équivalent

Posons $s = (M + m)/(M - m) > 1$. On établit une formule réelle pour $T_p(s)$. Définissons

$$z = s + \sqrt{s^2 - 1} > 1.$$

Comme $(s + \sqrt{s^2 - 1})(s - \sqrt{s^2 - 1}) = 1$, on a $z^{-1} = s - \sqrt{s^2 - 1}$ et $z + z^{-1} = 2s$.

La suite $F_p = (z^p + z^{-p})/2$ vérifie $F_0 = 1, F_1 = s$ et

$$2sF_p - F_{p-1} = \frac{(z + z^{-1})(z^p + z^{-p}) - (z^{p-1} + z^{-(p-1)})}{2} = \frac{z^{p+1} + z^{-(p+1)}}{2} = F_{p+1}.$$

Les mêmes valeurs initiales et la même relation de récurrence caractérisent $T_p(s)$; une récurrence donne donc $T_p(s) = F_p$.

La relation définissant les polynômes T_p donne également $T_p(-t) = (-1)^p T_p(t)$: c'est vrai pour $p = 0, 1$, et la relation de récurrence le conserve. Ainsi

$$|T_p(-s)| = T_p(s) = (z^p + z^{-p})/2.$$

Calculons z . Puisque $M > m > 0$,

$$\sqrt{s^2 - 1} = \frac{2\sqrt{Mm}}{M - m},$$

$$z = \frac{M + m + 2\sqrt{Mm}}{M - m} = \frac{\sqrt{M} + \sqrt{m}}{\sqrt{M} - \sqrt{m}}.$$

En posant $q = (\sqrt{M} - \sqrt{m})/(\sqrt{M} + \sqrt{m})$, on a $0 < q < 1$ et $z = 1/q$. Par conséquent

$$\left|T_p\left(\frac{m+M}{m-M}\right)\right| = \frac{q^{-p} + q^p}{2} = \frac{1}{2}q^{-p}(1 + q^{2p}).$$

Comme $q^{2p} \rightarrow 0$,

$$\left|T_p\left(\frac{m+M}{m-M}\right)\right| \underset{p \rightarrow +\infty}{\sim} \frac{1}{2}q^{-p}.$$

La borne optimisée possède même l'expression exacte

$$\frac{1}{|T_p(-s)|} = \frac{2q^p}{1 + q^{2p}} \underset{p \rightarrow +\infty}{\sim} 2q^p.$$

Rectification de l'énoncé original

Le PDF demandait l'équivalent $2^{p-1}((M+m)/(M-m))^p$, soit $(2s)^p/2$. Pour $s > 1$ fixé, cette expression n'est pas un équivalent de $T_p(s)$. En effet, l'expression exacte donne

$$\frac{|T_p(-s)|}{2^{p-1}s^p} = \left(\frac{s + \sqrt{s^2 - 1}}{2s} \right)^p (1 + (s + \sqrt{s^2 - 1})^{-2p}).$$

Comme $0 < (s + \sqrt{s^2 - 1})/(2s) < 1$, ce rapport tend vers 0, et non vers 1. Le coefficient dominant de T_p en tant que polynôme en t ne permet pas de déterminer sa valeur asymptotique lorsque son degré p varie à $t = s$ fixé. L'équivalent a donc été remplacé dans le sujet par la formule correcte établie ci-dessus.

Comparaison avec le pas constant optimal

Le facteur de la méthode constante optimale est $r = (M - m)/(M + m)$. En écrivant $M = a^2$, $m = b^2$ avec $a > b > 0$, on obtient

$$\frac{q}{r} = \frac{a^2 + b^2}{(a + b)^2} < 1,$$

car $(a + b)^2 = a^2 + b^2 + 2ab > a^2 + b^2$. Ainsi $0 < q < r < 1$.

La méthode à pas constant possède la borne $r^p \|E_0\|$; celle optimisée pour l'horizon p possède la borne $(2q^p/(1 + q^{2p})) \|E_0\|$, équivalente à $2q^p \|E_0\|$ lorsque $E_0 \neq 0$. Le rapport de leurs facteurs vérifie

$$\frac{2q^p}{(1 + q^{2p})r^p} \rightarrow 0.$$

La seconde borne décroît donc asymptotiquement plus vite. Elle est en outre au plus égale à la première pour chaque p , puisque $(1 - \alpha_* t)^p$ est un polynôme admissible du problème de minimisation et que Q_* minimise le supremum.

La comparaison porte sur des bornes uniformes pour les valeurs propres dans le segment $[m, M]$. L'erreur réellement observée dépend aussi de l'erreur initiale et du spectre effectif. Les pas ci-dessus dépendent de l'horizon p : en faisant varier p , on

compare les résultats optimisés après ce nombre d'étapes, et non les préfixes d'une unique suite de pas fixée à l'avance.

Cas $m = M$

Dans ce cas, $A = mI$. Le segment spectral se réduit à $\{m\}$ et

$$\sup_{m \leq t \leq M} |P_p(t)| = |P_p(m)|.$$

Pour tout $p \geq 1$, le minimum est 0 : il suffit qu'un des pas vaille $1/m$, ce qui annule un facteur $1 - \alpha_j m$. Par exemple, on prend $\alpha_0 = 1/m$ et les autres pas strictement positifs arbitraires. Alors $X_1 = X^*$ et tous les itérés suivants sont égaux à X^* .

La méthode constante optimale atteint déjà cette solution en une étape. Les expressions faisant intervenir $M - m$ au dénominateur ne s'appliquent pas ; aucune comparaison asymptotique supplémentaire n'est nécessaire.

Relevé des adaptations

I.1.d : la demande sur l'espace propre associé à 0 est remplacée par le calcul du noyau et la vérification que 0 est effectivement propre. Pour $n = 1$, le noyau est nul et 0 n'est pas une valeur propre.

II.1.c : les formules de racines avec division par p sont limitées à $p \geq 1$; le cas $p = 0$, où $T_0 = 1$, est précisé.

II.2 : l'hypothèse $p \geq 1$ est explicitée pour que les nœuds $\cos(j\pi/p)$ soient définis.

II.3 : le cas $p = 0$ est traité séparément ; l'unique polynôme admissible est la constante 1.

III.1.c : l'équivalence de convergence est rectifiée en précisant « pour tout vecteur initial ». Elle ne vaut pas pour chaque vecteur initial particulier indépendamment de ses composantes propres. Les cas $p = 0$ et $\lambda_1 = \lambda_n$ sont explicités.

III.2.c : les formules contenant $\lambda_n - \lambda_1$ au dénominateur sont réservées à $\lambda_1 < \lambda_n$; le cas d'égalité est traité sans division par zéro. L'équivalent incorrect

$2^{p-1}((\lambda_n + \lambda_1)/(\lambda_n - \lambda_1))^p$ est remplacé par $\frac{1}{2}q^{-p}$, avec

$q = (\sqrt{\lambda_n} - \sqrt{\lambda_1})/(\sqrt{\lambda_n} + \sqrt{\lambda_1})$. La dépendance du choix des pas à l'horizon p est précisée pour la comparaison finale.