Université Paris-Dauphine
Département MIDO
DUMI2E troisième année

Optimisation numérique
Gérard Lebourg
année 2008/2009

feuilles de travaux dirigés

Les notes biographiques sont pour partie tirées de WIKIPEDIA (`http://www.wikipedia.org/`).

Normes de matrices

Exercice 1 (rayon spectral). Pour toute matrice A réelle de taille $N \times N$, on note $\mathrm{sp}(A)$ le *spectre* de A, c'est-à-dire l'ensemble des valeurs propres (réelles ou complexes) de A, et $\rho(A)$ le *rayon spectral* de A, c'est-à-dire le plus petit réel majorant le module des valeurs propres de A. Montrer que :

1. $\rho(A^T) = \rho(A)$,
2. si A est symétrique (*i.e.*, $A^T = A$) alors $\mathrm{sp}(A)$ ne contient que des réels,
3. $\rho(A^2) = \rho(A)^2$,
   Indication : on pourra remarquer que $A^2 - \lambda^2 I = (A - \lambda I)(A + \lambda I)$.
4. si A est inversible alors $\lambda \in \mathrm{sp}(A^{-1}) \Leftrightarrow \lambda^{-1} \in \mathrm{sp}(A)$,
5. pour toute norme *matricielle* (*i.e.*, toute norme vérifiant $\|A\,B\| \le \|A\|\,\|B\|$) : $\rho(A) \le \|A\|$.

Exercice 2 (norme subordonnée à la norme euclidienne). Soit $\|\cdot\|_k$ une norme sur $\mathbb{R}^N$. On note encore $\|\cdot\|_k$ la norme matricielle qui lui est *subordonnée*, définie par :

$$\|A\|_k = \sup_{\|x\|_k \le 1} \|A\,x\|_k, \ \forall A \in \mathbb{R}^{N\times N}.$$

1. Prouver que l'on a toujours $\rho(A) \le \|A\|_k$.
2. Résoudre le problème consistant à maximiser ${\|A\,x\|_2}^2$ sous la contrainte ${\|x\|_2}^2 = \sum_{i=1}^N x_i^2 \le 1$ et en déduire que $\|A\|_2 = \rho(A^T A)^{1/2}$.
   Indication : on pourra remarquer que ${\|Ax\|_2}^2 = \langle x, A^T Ax\rangle$.
3. Comparer $\rho(A)$ et $\|A\|_2$ lorsque $A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$.
4. Vérifier que $\|A\|_\infty = \max\limits_{i\in\{1,\dots,N\}} \sum_{j=1}^N |A_{ij}|$ et en déduire que $\|A\|_1 = \max\limits_{j\in\{1,\dots,N\}} \sum_{i=1}^N |\,A_{ij}|$.
   Indication : on pourra remarquer que $\|x\|_1 = \sup\limits_{\|y\|_\infty \le 1} \langle x, y\rangle$ *et* $\|x\|_\infty = \sup\limits_{\|y\|_1 \le 1} \langle x, y\rangle$.

Exercice 3 (conditionnement). Soit A une matrice de taille $N \times N$ inversible. On note $\gamma_k(A) = \|A\|_k \|A^{-1}\|_k$ le *conditionnement* de la matrice A associé au choix de la norme $\|\cdot\|_k$ sur $\mathbb{R}^N$.

1. Si $A\,x = b$ et $A(x + \Delta x) = b + \Delta b$, où $A \in \mathbb{R}^{N\times N}$ et b un vecteur de $\mathbb{R}^N$, montrer que :

$$\frac{\|\Delta x\|_k}{\|x\|_k} \le \gamma_k(A)\,\frac{\|\Delta b\|_k}{\|b\|_k}.$$

   Comment interpréter cette inégalité ?

2. Avec le choix de $A = \begin{pmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{pmatrix}$, $x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ et $\Delta x = \begin{pmatrix} 82 \\ -136 \\ 35 \\ -21 \end{pmatrix}$, on trouve $b = A\,x = \begin{pmatrix} 32 \\ 23 \\ 33 \\ 31 \end{pmatrix}$, $\Delta b = A(\Delta x) = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$ et $\|\Delta x\|_2 \simeq 164$.
En déduire une minoration du conditionnement $\gamma_2(A)$.

Exercice 4 (norme de Frobenius [1]). Montrer que :

1. $\|A\| = \mathrm{tr}(A^T A)^{1/2} = \left(\sum_{i,j=1}^N |A_{ij}|^2\right)^{1/2}$ définit une norme sur $\mathbb{R}^{N\times N}$,
2. si P est une matrice orthogonale de taille $N \times N$ (*i.e.*, une matrice de taille $N \times N$ vérifiant $P^T P = I$), alors $\|P\,A\| = \|A\,P\| = \|A\|$ et donc $\|P^T A\,P\| = \|A\|$,
3. $\|\cdot\|$ est une norme matricielle, mais elle n'est subordonnée à aucune norme sur $\mathbb{R}^N$.
Indication : on pourra remarquer que $\|I\| = \sqrt{N}$.

Formes quadratiques et ellipticité

Exercice 5 (encadrement des formes quadratiques). Une *forme quadratique* q sur $\mathbb{R}^N$ est une application de la forme
$$q(x) = \langle A\,x, x\rangle\,,\ \forall x \in \mathbb{R}^N,$$
où A est une matrice réelle symétrique de taille $N \times N$.

1. Calculer $\nabla q(x)$ puis résoudre le problème consistant à minimiser $q(x)$ sous la contrainte ${\|x\|_2}^2 = 1$.
2. En déduire l'encadrement : $\lambda_{min}{\|x\|_2}^2 \le q(x) \le \lambda_{max}{\|x\|_2}^2$, (*i.e.*, $\lambda_{min} I \le A \le \lambda_{max} I$).

Exercice 6 (matrices définies positives, semi-définies positives et critère de Sylvester [2]). Une matrice A réelle symétrique de taille $N \times N$ est dite *semi-définie positive* [3] si la forme quadratique qui lui est associée $q(x) = \langle A\,x, x\rangle$ est une fonction positive. Elle est dite *définie positive* si en outre $q(x) = 0 \Rightarrow x = 0$.

1. Vérifier que A semi-définie positive (resp. définie positive) équivaut à $\mathrm{sp}(A) \subset [0, +\infty[$ (resp. $\mathrm{sp}(A) \subset]0, +\infty[$).
2. Soient A une matrice définie positive de taille $N \times N$, v un vecteur de $\mathbb{R}^N$, a un réel strictement positif, $w = A^{-1}v$ et B la matrice de taille $(N+1)\times(N+1)$ s'écrivant sous forme de blocs $B = \begin{pmatrix} A & v \\ v^T & a \end{pmatrix}$. On suppose que $\det(B) > 0$.
   (a) Vérifier que
$$\det(B) = \begin{vmatrix} A & v \\ v^T & a \end{vmatrix} = \begin{vmatrix} A & v - A\,w \\ v^T & a - v^T w \end{vmatrix} = (a - \langle w, A\,w\rangle)\det(A).$$

1. Ferdinand Georg Frobenius (26 octobre 1849 - 3 août 1917) était un mathématicien allemand. Il s'intéressa principalement à la théorie des groupes et à l'algèbre linéaire, mais travailla également en analyse et en théorie des nombres.

2. James Joseph Sylvester (3 septembre 1814 - 13 mars 1897) était un mathématicien et géomètre anglais. Il travailla sur les formes algébriques, particulièrement sur les formes quadratiques et leurs invariants, et la théorie des déterminants. On lui doit l'introduction de nombreux objets, notions et notations mathématiques, comme le *discriminant* ou la *fonction indicatrice d'Euler*.

3. On dit aussi simplement qu'elle est *positive*.

(b) En déduire que $\begin{pmatrix} u^T & \alpha \end{pmatrix} \begin{pmatrix} A & v \\ v^T & a \end{pmatrix} \begin{pmatrix} u \\ \alpha \end{pmatrix} \geq \langle u + \alpha\, w, A(u + \alpha\, w)\rangle$, pour tous $u \in \mathbb{R}^N$ et $\alpha \in \mathbb{R}$.

(c) En conclure que la matrice B est définie positive.

3. Réciproquement, montrer que $B = \begin{pmatrix} A & v \\ v^T & a \end{pmatrix}$ est définie positive implique que $\det(B) > 0$, que A est définie positive et que $a > 0$.
4. En déduire une règle simple (appelée *critère de Sylvester*) permettant de vérifier le caractère défini positif d'une matrice symétrique donnée.
5. **Application :** utiliser cette règle pour vérifier que la matrice $A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$ est définie positive.

Exercice 7 (solution au sens des moindres carrés d'un système linéaire). Une *fonction quadratique* F sur $\mathbb{R}^N$ est une application de la forme

$$F(x) = \frac{1}{2}\,\langle A\,x, x\rangle - \langle b, x\rangle + c,\ \forall x \in \mathbb{R}^N,$$

où A est une matrice réelle symétrique de taille $N \times N$, b un vecteur de $\mathbb{R}^N$ et $c \in \mathbb{R}$. Si A est une matrice réelle de taille $M \times N$ et b un vecteur de $\mathbb{R}^M$, la fonction $F(x) = \|A\,x - b\|^2$ définit une fonction quadratique sur $\mathbb{R}^N$.

1. Calculer $\nabla F(x)$ et $\nabla^2 F(x)$.
2. Si la matrice A est de rang égal à N, montrer que la fonction F atteint son minimum en un unique point $x^\star$, solution du système des *équations normales*[4] $A^T A x = A^T b$.
3. Déterminer A, b et $x^\star$ dans le cas où $F(x_1, x_2) = (x_1+x_2-2)^2+(2\,x_1+x_2-3)^2+(3\,x_1+x_2-5)^2$.

Exercice 8 (ellipticité). Une fonction numérique F de classe $\mathscr{C}^2$ sur $\mathbb{R}^N$ est dite *elliptique* s'il existe une constante strictement positive c – appelée constante d'ellipticité de F – telle que : $\big\langle v, \nabla^2 F(x)\, v\big\rangle \geq c\,\|v\|^2$ pour tous x et v dans $\mathbb{R}^N$.

1. Montrer que toute fonction elliptique est strictement convexe.
2. Si c est une constante d'ellipticité pour F alors $F(x) \geq F(x_0)+\langle \nabla F(x_0), x - x_0\rangle+\dfrac{c}{2}\,\|x-x_0\|^2$ pour tout couple de points x_0 et x dans $\mathbb{R}^N$. Interpréter.
3. En déduire que toute fonction elliptique est coercive (*i.e.*, $\lim\limits_{\|x\|\to+\infty} F(x) = +\infty$) et que ses ensembles de niveau sont des convexes compacts.

Exercice 9 (fonctions quadratiques elliptiques). Vérifier que :

1. l'application $F(x) = \frac{1}{2}\,\langle A\,x, x\rangle - \langle b, x\rangle + c$ définit une fonction quadratique elliptique sur $\mathbb{R}^N$ si et seulement si la matrice A symétrique d'ordre N est définie positive,
2. si A est une matrice réelle de taille $M \times N$ et b est un vecteur de $\mathbb{R}^M$ alors $F(x) = \|A\,x - b\|^2$ définit une fonction quadratique elliptique sur $\mathbb{R}^N$ si et seulement si $\mathrm{rg}(A) = N$,
3. l'application $F(x) = (x_N - 1)^2 + x_0^2 + \sum_{i=0}^{N-1}(x_{i+1} - x_i)^2$ définit une fonction elliptique sur $\mathbb{R}^{N+1}$,
4. les deux fonctions F et G définies sur $\mathbb{R}^2$ par

$$F(x, y) = x^2 + 2\,y^2 + 3\,x\,y \text{ et } G(x, y) = 2\,x^2 + 3\,y^2 + 4\,x\,y,$$

sont quadratiques, mais que $\inf G$ est fini alors que $\inf F = -\infty$. Donner une explication.

4. Le vecteur $x^\star$ est alors appelé la solution *au sens des moindres carrés* du système $A\,x = b$.

Utilisation des conditions de Karush [5]-Kuhn [6]-Tucker [7]

Exercice 10. Un établissement de restauration rapide sert trois sortes de plats :

- des hamburgers pour 3 euros,
- du poulet-frites pour 6 euros,
- des sushis pour 9 euros.

On connait le chiffre d'affaires et le nombre de plats vendus chaque jour. À partir de ces données, on a pu estimer le prix moyen du repas à 5.25 euros. Un client entre dans le restaurant. Quelle est la probabilité qu'il commande un poulet-frites ?
Indication : on pourra utiliser l'estimateur du maximum d'entropie qui retourne l'unique solution du problème

$$\min_{\substack{p+q+r=1 \\ 3p+6q+9r=5.25 \\ p,q,r\geq 0}} (p \ln p + q \ln q + r \ln r)$$

pour estimer la probabilité cherchée.

Exercice 11. Le problème du colis postal consiste à maximiser le volume d'une boîte parallélépipédique de dimensions x_1, x_2 et x_3 sous les contraintes $0 \leq x_i \leq 42$, $i = 1, 2, 3$, et $x_1 + 2\,x_2 + 2\,x_3 \leq 72$, ce que l'on résume dans le problème d'optimisation suivant :

$$\max_{\substack{0\leq x_i \leq 42,\ i=1,2,3, \\ x_1+2\,x_2+2\,x_3\leq 72.}} x_1\, x_2\, x_3.$$

1. Prouver que le problème ci-dessus admet une solution au moins.
2. Vérifier qu'il équivaut au problème convexe sous contraintes linéaires suivant :

$$\min_{\substack{0< x_i \leq 42,\ i=1,2,3, \\ x_1+2\,x_2+2\,x_3\leq 72.}} (-\ln x_1 - \ln x_2 - \ln x_3),$$

   et en déduire qu'il admet une unique solution $(\hat{x_1}, \hat{x_2}, \hat{x_3})$.
3. Écrire les conditions KKT associées au problème, puis prouver que la contrainte $x_1 + 2\,x_2 + 2\,x_3 \leq 72$ est saturée en $(\hat{x_1}, \hat{x_2}, \hat{x_3})$ et que $\hat{x_2} = \hat{x_3}$. Résoudre finalement le problème du colis postal.

Exercice 12. Le problème de la « boîte » consiste à minimiser la surface des parois d'une boîte parallélépipédique de dimensions x, y et z sous la contrainte que le volume contenu dans la boîte soit égal une valeur prescrite $V > 0$, c'est-à-dire :

$$\min_{\substack{x\,y\,z=V \\ x,y,z>0}} (x\,y + 2\,x\,z + 2\,y\,z).$$

En utilisant le changement de variables : $u = \ln x$, $v = \ln y$ et $w = \ln z$, récrire ce problème comme un problème convexe sous contraintes linéaires. Prouver ensuite l'existence d'une unique solution et la calculer.

5. William Karush (1er mars 1917 - 22 février 1997) était un mathématicien américain. Il fût le premier à établir les conditions nécessaires pour résoudre des problèmes d'optimisation sous contraintes d'inégalité non-linéaires dans sa thèse de maîtrise soutenue en 1939. Il ne devint cependant célèbre qu'après la publication en 1951 d'un acte de conférence de Harold W. Kuhn et Albert W. Tucker réintroduisant les conditions qui portent aujourd'hui leurs trois noms.

6. Harold William Kuhn (né le 29 juillet 1925) est un mathématicien et économiste américain qui travailla principalement en théorie des jeux.

7. Albert William Tucker (28 novembre 1905 - 25 janvier 1995) était un mathématicien américain d'origine canadienne qui apporta d'importantes contributions en topologie, théorie des jeux et programmation non-linéaire.

Exercice 13. On considère le problème :

$$(\mathcal{P}) \qquad \max_{\substack{x\geq 0\\ \sum_{i=1}^n x_i=1}} \sum_{i=1}^{n} \ln(x_i+a_i),$$

où $a=(a_1,\ldots,a_n)$ est un vecteur donné de $\mathbb{R}^n$ dont toutes les composantes sont strictement positives et $x=(x_1,\ldots,x_n)$ est le vecteur des variables du problème.

1. Prouver l'existence d'une unique solution $\hat{x}$ de $(\mathcal{P})$.
2. Écrire les conditions KKT associées au problème $(\mathcal{P})$.
3. Prouver l'existence d'un réel ξ tel que $\hat{x_i}=\max(0,\xi-a_i)$, $i=1,\ldots,n$.
4. On pose $n=4$ et $a=(0.25,0.5,0.75,1)$. Trouver $\hat{x}$.

Exercice 14. On considère le problème :

$$(\mathcal{P}) \quad \min_{A\,x\leq b} \frac{1}{2}\,\|x-a\|^2,$$

où les vecteurs $a\in\mathbb{R}^N$, $b\in\mathbb{R}^M$ et la matrice $A\in\mathbb{R}^{M\times N}$ sont donnés.

1. Écrire les conditions KKT associées au problème $(\mathcal{P})$.
2. En déduire que le vecteur $M=(\mu_1,\ldots,\mu_M)$ des multiplicateurs associés aux contraintes du problème $(\mathcal{P})$ est solution du problème :

$$(\mathcal{P}^\star) \quad \min_{M\leq 0} \frac{1}{2}\left\|A^T M\right\|^2+\langle M, A\,a-b\rangle\,.$$

3. Prouver l'unicité de la solution de $(\mathcal{P}^\star)$ lorsque la matrice A est de rang M.
4. Expliciter $(\mathcal{P}^\star)$ dans le cas où $a=\begin{pmatrix}1 & 2.4 & 3 & 1.7 & 0.9\end{pmatrix}^T$, $A=\begin{pmatrix}1 & 1 & 1 & 1 & 1\end{pmatrix}$ et $b=1$. Le résoudre et en déduire l'unique solution de $(\mathcal{P})$.

Méthodes de gradient

Exercice 15. On considère le problème de Toricelli [8] :

$$\min_{x\in\mathbb{R}^2} \left(\|x-a\|+\|x-b\|+\|x-c\|\right),$$

avec $a=(1,4)$, $b=(-3,-4)$ et $c=(5,1)$.

1. Faire un dessin et proposer une solution approchée raisonnable x_0.
2. Trouver une direction de descente au point proposé x_0 et en conclure que ce n'est pas un minimiseur local.

Exercice 16. On considère sur $\mathbb{R}^2$ la fonction $F(x,y)=x^2+2\,y^4-3\,x\,y^2$.

1. Vérifier que $(0,0)$ est un point critique dégénéré [9] de la fonction F.
2. Montrer que, pour toute direction (u,v) dans $\mathbb{R}^2$, la fonction d'une variable $\varphi(t)=F(t\,u,t\,v)$ reste strictement positive pour $t>0$ suffisamment petit. En conclure que F n'admet aucune direction de descente en $(0,0)$.

8. Ce problème, dû au physicien et mathématicien italien Evangelista Torricelli (1608–1647), s'énonce de la manière suivante : *« soit un triangle ABC dont les mesures des angles sont inféreures à 120 degrés. Où doit se trouver le point M minimisant la somme $MA+MB+MC$? »*. La solution de ce problème est appelée le *point de Fermat* du triangle ABC.

9. On rappelle qu'un point critique d'une fonction est dit *dégénéré* lorsqu'il annule le déterminant de la matrice hessienne de cette fonction.

3. Développer $(x-y^2)(x-2\,y^2)$ et en déduire que $(0,0)$ n'est pas un minimiseur local de F.
4. Expliquer pourquoi une telle situation ne peut se produire qu'en un point critique dégénéré.

Exercice 17. On considère l'algorithme du gradient à pas fixe appliqué à la minimisation de $F(x) = x^4$.

1. Écrire l'équation de récurrence définissant l'algorithme.
2. Vérifier que $x \in S_0 = \{x \in \mathbb{R} \,|\, x^4 \leq x_0^4\}$ implique que $\nabla^2 F(x) \leq K\,I$, où $K = 12\,x_0^2$.
3. Le théorème énoncé en cours garantit la convergence de l'algorithme du gradient à pas fixe pour $\rho < \frac{1}{6\,x_0^2}$. Vérifier cependant que la convergence n'est jamais linéaire et expliquer pourquoi.

Exercice 18. On souhaite calculer une valeur approchée de $x^\star = \sqrt[3]{2}$ en utilisant l'algorithme du gradient à pas fixe pour minimiser la fonction d'une variable $F(x) = x^4 - 8\,x + 1$.

1. Écrire l'équation de récurrence définissant l'algorithme.
2. L'initialisation étant fixée égale à $x_0 = 1$, comment choisir le pas pour assurer la convergence de l'algorithme ?
3. Vérifier par un calcul direct que cette convergence est alors linéaire, et qu'elle devient superlinéaire pour une valeur du pas égale $\frac{1}{12\,\sqrt[3]{4}}$.

Exercice 19. On considère la fonction $F(x,y) = x^4 + y^4 + 4\,x\,y$.

1. Prouver que tous ses ensembles de niveau sont compacts.
2. Trouver tous ses points critiques et la nature de chacun d'entre eux.
3. On applique l'algorithme du gradient à pas fixe à la minimisation de F avec un pas fixé égal à ρ. Prouver que, pour toute initialisation (x_0, y_0), l'algorithme converge, pour ρ assez petit, vers un point critique de F.
4. On choisit $x_0 = y_0 = 1$. Que se passe-t-il si $\rho = 0.25$? Si $\rho = 0.5$?

Exercice 20. On considère l'algorithme du gradient à pas fixe appliqué à la minimisation de la fonction $F(x,y) = x^2 + 2\,y^2$.

1. Écrire explicitement les équations de récurrence définissant l'algorithme. Pour quelles valeurs du pas de descente est-il convergent ?
2. Vérifier que le rayon spectral de la matrice $I - \rho\,H$, avec ρ le pas de descente et H la matrice hessienne (constante) de F, est minimal pour $\rho = 1/3$.
3. On fixe $\rho = 1/3$ et on choisit l'initialisation $(x_0, y_0) = (2,1)$. Calculer les premiers itérés et décrire alors le comportement de l'algorithme.
4. Quels seraient les itérés calculés par l'algorithme du gradient à pas optimal à partir la même initialisation ?

Exercice 21. Soit A une matrice symétrique d'ordre N définie positive.

1. Vérifier que $F(x) = \frac{1}{2}\,\langle x, A\,x\rangle - \langle b, x\rangle$ atteint son minimum en un unique point $x^\star$, solution du système linéaire $A\,x = b$.
2. Expliciter les relations de récurrence permettant de calculer $x^\star$ en utilisant :

   a. l'algorithme du gradient à pas fixe,

   b. l'algorithme du gradient à pas optimal,

   tous deux appliqués à la minimisation de F.

Exercice 22 (inégalité de Kantorovitch[10] et convergence de l'algorithme du gradient à pas optimal). On considère l'algorithme du gradient à pas optimal appliqué à la minimisation de la fonction quadratique $F(x) = \frac{1}{2}\langle x, A\,x\rangle - \langle b, x\rangle$, où A est une matrice symétrique d'ordre N définie positive et b un vecteur de $\mathbb{R}^N$, tous deux donnés.

1. Si x_k est le $k^{\text{ème}}$ point calculé par l'algorithme, déterminer explicitement le pas optimal au point x_k dans la direction opposée au gradient $g_k = \nabla F(x_k)$ en fonction de A et de g_k.
2. Vérifier que $\|x\|_A = \langle A\,x, x\rangle^{1/2}$ définit une norme sur $\mathbb{R}^N$ et établir l'identité :
$$\|x_{k+1} - x^\star\|_A{}^2 = \left(1 - \frac{\|g_k\|_2{}^4}{\langle A^{-1}g_k, g_k\rangle\ \langle A\,g_k, g_k\rangle}\right)\|x_k - x^\star\|_A{}^2,$$
où $x^\star$ est l'unique minimiseur de F (on pourra remarquer que $g_k = A\,(x_k - x^\star)$).
3. En admettant alors l'*inégalité de Kantorovitch* :
$$\langle x, A\,x\rangle\,\langle x, A^{-1}x\rangle \leq \frac{(1+\gamma)^2}{4\,\gamma}\|x\|_2{}^4,$$
où γ est le conditionnement de la matrice A, obtenir :
$$\frac{\|x_{k+1} - x^\star\|_A}{\|x_k - x^\star\|_A} \leq \frac{\gamma - 1}{\gamma + 1}.$$
4. **(question difficile)** Démontrer l'inégalité de Kantorovitch pour toute matrice symétrique d'ordre N définie positive[11].

Exercice 23. Les N premiers moments $M_k = \int_0^1 t^k\,f(t)\,\mathrm{d}t$, $1 \leq k \leq N$, d'une variable aléatoire réelle X à valeurs dans $[0, 1]$ étant connus, on cherche à estimer la densité inconnue f de X par un polynôme P de degré N minimisant l'écart quadratique $\frac{1}{2}\int_0^1 |P(t) - f(t)|^2\,\mathrm{d}t$.

1. Prouver que ce problème admet toujours une unique solution.
2. En posant $P(t) = \sum_{i=0}^N x_i\,t^i$, montrer qu'il équivaut à minimiser sur $\mathbb{R}^{N+1}$ la fonction F définie par $F(x) = \frac{1}{2}\langle H\,x, x\rangle - \langle M, x\rangle$, où H est la matrice de Hilbert[12] d'ordre $N+1$, de terme général $H_i^j = 1/(i+j-1)$ et $M = (1, M_1, \cdots, M_N)^T$, puis que H est définie positive.
3. Pour résoudre ce problème, on programme sous MATLAB une version adaptée de l'algorithme du gradient à pas optimal :

```
function sol=gradopt(M,itermax)
N=length(M); M=[1 M]; iternum=0;
sol=[1 zeros(1,N)];
N=N+1; H=hilb(N);
g=sol*H-M; d=-g;
while iternum < itermax
   iternum=iternum+1;
   rho=-(g*d')/(d*H*d');
   sol=sol+rho*d;
   g=sol*H-M; d=-g;
end
```

10. Leonid Vitalievitch Kantorovitch (Леони́д Вита́льевич Канторо́вич, 19 janvier 1912 - 7 avril 1986) était un mathématicien et économiste russe (soviétique), connu pour ses contributions à la théorie de l'allocation optimale des ressources. Il reçut le « Prix Nobel » d'économie en 1975.

11. Il suffira de démontrer l'inégalité dans le cas où A est diagonale (on expliquera pourquoi). Dans ce cas, le problème se réduit à maximiser la fonction $\left(\sum_{i=1}^N \Lambda_i {x_i}^2\right)\left(\sum_{i=1}^N {\Lambda_i}^{-1} {x_i}^2\right)$ des variables réelles x_i, $i = 1, \ldots, N$, sous la contrainte $\sum_{i=1}^N {x_i}^2 = 1$, où les réels Λ_i, $i = 1, \ldots, N$, sont les valeurs propres strictement positives de A. En écrivant les conditions nécessaires d'optimalité pour ce problème, on montrera que deux au plus des x_i sont non nuls.

12. David Hilbert (23 janvier 1862 - 14 février 1943) était un mathématicien allemand, souvent considéré comme un des plus grands mathématiciens du XX$^{\text{ème}}$ siècle. Il a créé ou développé un large éventail d'idées fondamentales, que ce soit la théorie des invariants, l'axiomatisation de la géométrie ou les fondements de l'analyse fonctionnelle.

Commenter le choix de l'initialisation et le calcul du pas optimal.

4. On teste cette procédure dans le cas élémentaire $N = 1$ et $M_1 = 1/3$.
   (a) Vérifier par un calcul explicite que la solution du problème est alors $(2, -2)$.
   (b) Il faut pourtant `itermax`$\geq$ 50 pour obtenir une approximation de cette solution à 10^{-3} près. Expliquer ce médiocre comportement de l'algorithme en calculant une valeur approchée du conditionnement de la matrice de Hibert d'ordre deux

$$H = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}.$$

5. Comment modifier le code de la procédure ci-dessus pour obtenir une version adaptée de l'algorithme du gradient conjugué ? Calculer alors explicitement les deux premiers itérés et vérifier la finie convergence de l'algorithme.

Exercice 24 (exercice difficile). Soit $F : \mathbb{R} \to \mathbb{R}$ une fonction analytique (*i.e.*, une fonction $\mathscr{C}^\infty$ qui est la somme de sa série de Taylor[13] en tout point) et $(x_k)_k$ la suite des points calculés par l'algorithme du gradient à pas fixe appliqué à la minimisation de F. On suppose que $(x_k)_k$ tend vers $x^\star$ et que :
- $(x_k)_k$ n'est pas *stationnaire*, c'est-à-dire qu'il n'existe pas d'indice n tel que la suite soit constante à partir du rang n,
- $(x_k)_k$ *oscille* autour de $x^\star$, c'est-à-dire que, pour tout indice k, on peut trouver $k_g \geq k$ et $k_d \geq k$ tels que $x_{k_g} \leq x^\star \leq x_{k_d}$.

Prouver alors que $x^\star$ est un minimum local de f.

Méthodes newtoniennes

Exercice 25. On souhaite calculer une valeur approchée de la racine $n^{\text{ème}}$ d'une réel $a \geq 0$ donné en minimisant $F(x) = x^{n+1} - a\,(n+1)\,x$ par la méthode de Newton[14] à pas fixe avec $\rho = 1$.

1. Écrire l'équation de récurrence définissant l'algorithme.
2. Vérifier que celui-ci converge pour toute initialisation et que la convergence est quadratique si $a > 0$, mais seulement linéaire si $a = 0$. Donner une explication.

Exercice 26. On souhaite tester l'algorithme de Newton à pas fixe avec $\rho = 1$ appliqué à la minimisation de $F(x) = \ln(\cosh x)$.

1. Vérifier que :
   a. F est strictement convexe et coercive sur $\mathbb{R}$,
   b. $F(x) \leq F(x_0) \Rightarrow 0 < \dfrac{1}{\cosh^2 x_0} \leq F''(x) \leq 1$.
2. L'expérimentation numérique montre cependant que l'algorithme, initialisé avec $x_0 = 1.1$ diverge. Expliquer pourquoi.

Exercice 27 (*logistic regression*). Le modèle `logit` suppose la probabilité qu'un « individu » – en fait un triplet x listant ses mensurations : taille, poids et pointure – soit une fille égale à $P = (1 + \exp(\langle p, x\rangle + q))^{-1}$.

1. Vérifier que la probabilité que l'individu x soit un garçon est égale à $(1 + \exp(-\langle p, x\rangle - q))^{-1}$.

13. Brook Taylor (18 août 1685 - 30 novembre 1731) était un mathématicien, artiste peintre et musicien anglais. Il inventa le calcul aux différences finies et découvrit l'intégration par parties.

14. Sir Isaac Newton (4 janvier 1643 - 31 mars 1727) était un philosophe, mathématicien, physicien et astronome anglais. Figure emblématique des sciences, il est surtout reconnu pour sa théorie de la gravitation universelle et la création du calcul infinitésimal.

2. On cherche à estimer les paramètres p et q du modèle à partir d'un échantillon de N individus dont la classe, *i.e.*, le sexe, est connue en minimisant la log-vraisemblance

$$F(p,q) = \sum_{i=1}^{N} \ln\left(1 + \exp\left[\xi_i(\langle p, x_i\rangle + q)\right]\right)$$

de l'échantillon (pour tout indice i, $\xi_i = \pm 1$ est un « marqueur » de la classe « fille » ou « garçon » du $i^{\text{ème}}$ individu x_i). Vérifier que :

a. $\nabla^2 F(p,q) = \dfrac{1}{4} \displaystyle\sum_{i=1}^{N} \frac{1}{\cosh^2 \alpha_i} \begin{pmatrix} x_i\, x_i^T & x_i \\ x_i^T & 1 \end{pmatrix}$, avec $\alpha_i = \dfrac{\xi_i}{2}\left(\langle p, x_i\rangle + q\right)$,

b. la fonction F est strictement convexe sur $\mathbb{R}^4$ dès qu'il n'existe aucun hyperplan de $\mathbb{R}^3$ contenant les N points x_i de l'échantillon,

c. la fonction F est coercive sur $\mathbb{R}^4$ s'il n'existe aucun hyperplan de $\mathbb{R}^3$ « séparant » les classes, *i.e.*, tel que, pour tout indice i, $\xi_i\left(\langle p, x_i\rangle + q\right) \leq 0$.

3. En conclure que l'algorithme de Newton appliqué à la minimisation de F converge, pour toute initialisation (p_0, q_0), vers l'unique minimiseur de F lorsqu'il n'existe aucun hyperplan de $\mathbb{R}^3$ séparant les classes.

Exercice 28. Soient F une fonction de classe $\mathscr{C}^2$, strictement convexe et coercive sur $\mathbb{R}$, x_0 un point de $\mathbb{R}$ et $L > 0$ une constante donnés. On suppose que la dérivée seconde de F ne s'annule jamais et qu'elle est L-Lipschitzienne sur l'ensemble $S_0 = \{x \in \mathbb{R} \,|\, F(x) \leq F(x_0)\}$.

1. On considère l'algorithme défini par la relation de récurrence :

$$x_{k+1} = \arg\min \left(F(x) + \frac{L}{6}\,|x - x_k|^3 \right).$$

Prouver successivement que :

a. pour tout entier k, x_{k+1} est bien défini et de manière unique,

b. la suite $(F(x_k))_k$ est décroissante et la différence $x_{k+1} - x_k$ tend vers 0,

c. la suite $(F'(x_k))_k$ tend vers 0,

d. la suite $(x_k)_k$ converge vers l'unique minimiseur $x^\star$ de F,

e. il existe une constante $c > 0$ telle que :

$$F(x^\star) + \frac{c}{2}\,|x_{k+1} - x^\star|^2 \leq F(x_{k+1}) \leq F(x^\star) + \frac{L}{6}\,|x^\star - x_k|^3,$$

et donc la convergence est superlinéaire.

2. Connaissant $y_0 = x_k$, on décide de calculer une valeur approchée de $y^\star = x_{k+1}$ en minimisant $G(y) = F(y) + \frac{L}{6}\,|y - y_0|^3$ au moyen d'une variante de la méthode de Newton à pas fixe et égal à l'unité, obtenue en ajoutant, à chaque itération de l'algorithme, l'étape suivante :

$$\texttt{tant que } G(y_{l+1}) > G(y_l),\ y_{l+1} = \frac{y_{l+1} + y_l}{2}.$$

a. Établir l'implication

$$y_{l+1} = y_l - \frac{G'(y_l)}{G''(y_l)} \Rightarrow y_{l+1} - y^\star = (y_l - y^\star)\,\frac{G''(y_l) - G''(z_l)}{G''(y_l)},$$

où z_l est strictement compris entre y_l et $y^\star$.

b. Vérifier que G'' est décroissante sur $]-\infty, y_0[\,\cap S_0$ et croissante sur $]y_0, +\infty[\,\cap S_0$.

c. En conclure que la suite $(y_l)_l$ des points calculés par l'algorithme est monotone à partir d'un certain rang, qu'elle converge vers $y^\star$ et que la convergence est toujours quadratique.

3. **Application :** une variable aléatoire X peut prendre N valeurs distinctes $x_0 < \ldots < x_N$ connues. On cherche la distribution p_i de X d'entropie maximale et de moyenne m donnée :

$$(\mathcal{P}) \qquad \min_{\substack{\sum_{i=1}^N p_i = 1 \\ \sum_{i=1}^N p_i\, x_i = m}} \sum_{i=1}^N p_i \,\ln p_i .$$

a. Montrer que $p_i^\star = \dfrac{\exp(\mu^\star\, x_i)}{\sum_{i=1}^N \exp(\mu^\star\, x_i)}$ est l'unique solution de $(\mathcal{P})$ si

$$\mu^\star = \arg\min \sum_{i=1}^N \exp[\mu\,(x_i - m)]$$

(on pourra remarquer que $\mu^\star$ est le multiplicateur de Lagrange [15] associé à la contrainte $\sum_{i=1}^N p_i\, x_i = m$).

b. Montrer que la fonction $F(\mu) = \sum_{i=1}^N \exp(\mu(x_i - m))$ vérifie les hypothèses de l'exercice et suggérer un algorithme de calcul de $\mu^\star$. Comment choisir L si l'algorithme est initialisé avec $\mu_0 = 0$?

Exercice 29. On cherche à calculer la projection $\pi(a)$ d'un point donné a de $\mathbb{R}^N$ sur le polyèdre

$$C = \left\{x \in \mathbb{R}^N \mid A\,x \le b\right\},$$

où A est une matrice de taille $M \times N$ et de rang M et b un vecteur de $\mathbb{R}^M$ donnés.

1. Montrer que $\pi(a) = a - A^T\mu$, où $\mu \in \mathbb{R}^M$ minimise $F(\mu) = \frac{1}{2}\|A^T\mu\|^2 - \langle \mu, A\,a - b\rangle$ sous la contrainte $\mu \ge 0$.
2. On se propose alors de calculer une valeur approchée de μ en minimisant la fonction

$$F_\alpha(\mu) = F(\mu) - \alpha \sum_{i=1}^N \ln \mu_i$$

où $\alpha > 0$ est un paramètre destiné à tendre vers zéro.

a. Vérifier que F_α est elliptique sur $\mathbb{R}^M$.

b. En conclure que l'algorithme de Newton appliqué à la minimisation de F_α converge pour toute initialisation $\mu_0 > 0$ et que la convergence est quadratique.

Exercice 30 (formule de mise-à-jour de Broyden, Fletcher, Goldfarb et Shanno). Vérifier par un calcul direct que, si A_k est une matrice symétrique de taille $N \times N$ définie positive et que v_k et w_k sont deux vecteurs non nuls et non orthogonaux de $\mathbb{R}^N$, l'inverse de la matrice $A_k + \dfrac{w_k w_k^T}{w_k^T v_k} - \dfrac{A_k v_k v_k^T A_k}{v_k^T A_k v_k}$ est la matrice [16]

$$\left(I - \frac{v_k w_k^T}{w_k^T v_k}\right) B_k \left(I - \frac{w_k v_k^T}{w_k^T v_k}\right) + \frac{v_k v_k^T}{w_k^T v_k},$$

15. Joseph Louis Lagrange (Giuseppe Lodovico Lagrangia en italien, 25 janvier 1736 - 10 avril 1813) était un mathématicien et astronome franco-italien. Fondateur du calcul des variations avec Euler, il a également produit d'importantes contributions tant en analyse, qu'en géométrie, en théorie des groupes et en mécanique.

16. On pourra également vérifier, au prix d'un calcul laborieux mais mécanique, que cette matrice est également donnée par la formule

$$B_k + \frac{1}{w_k^T v_k}\left(1 + \frac{w_k^T B_k w_k}{w_k^T v_k}\right) v_k v_k^T - \frac{B_k w_k v_k^T + v_k w_k^T B_k}{w_k^T v_k}.$$

où B_k est l'inverse de A_k.

Exercice 31 (formule de mise-à-jour PSB). Soient v_k et w_k deux vecteurs de $\mathbb{R}^N$ et M_k une matrice symétrique de taille $N \times N$. On considère le problème en la matrice $M \in \mathbb{R}^{N\times N}$ suivant

$$(\mathcal{P}) \qquad \min_{\substack{v_k = M\, w_k \\ M = M^T}} \|M - M_k\|,$$

où $\|\cdot\|$ est la norme de Frobenius. Montrer que

1. si $w_k \neq 0$, alors le problème $(\mathcal{P})$ a une solution unique M_{k+1} et que celle-ci s'écrit [17]
$$M_{k+1} = M_k + \frac{(v_k - M_k w_k)w_k^T + w_k(v_k - M_k w_k)^T}{w_k^T w_k} - \frac{(v_k - M_k w_k)^T w_k}{(w_k^T w_k)^2} w_k w_k^T,$$
2. si $w_k = 0$ et $v_k \neq 0$ [18], alors le problème $(\mathcal{P})$ n'a pas de solution,
3. si $w_k = v_k = 0$ [19], alors le problème $(\mathcal{P})$ a comme unique solution $M_{k+1} = M_k$.

Exercice 32. On teste de l'algorithme quasi-Newton DFP appliqué à la minimisation de la fonction quadratique elliptique
$$F(x,y) = (x-1)^2 + (y-1)^2 + (x+y-1)^2$$
L'algorithme est initialisé avec le point $(x_0, y_0) = (0,0)$ et l'approximation de l'inverse de la matrice hessienne par $B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

1. Vérifier la finie convergence de l'algorithme.
2. Vérifier que l'algorithme calcule bien l'inverse de la matrice hessienne de F.
3. Qu'obtient-on avec la version BFGS de l'actualisation de B_k ?

Problèmes quadratiques

Exercice 33.

1. Vérifier que la matrice $A = \begin{pmatrix} 2 & 1 & 1 & 5 \\ -1 & 1 & 3 & 2 \end{pmatrix}$ est de rang égal à deux.
2. En déduire que AA^T est inversible et calculer $\left(AA^T\right)^{-1}$.
3. En déduire la solution de norme minimale du système d'équations
$$\begin{cases} 2\,x + y + z + 5\,t = 8, \\ -x + y + 3\,z + 2\,t = 0. \end{cases}$$

Exercice 34. Soient A une matrice de taille $N \times N$ et un vecteur $b \in \mathbb{R}^N$ donnés. On considère le problème
$$(\mathcal{P}) \qquad \min_{\substack{x \in \mathbb{R}^N \\ A\,x = 0}} \left(\frac{1}{2}\|x\|^2 - \langle b, x\rangle\right).$$

1. Déterminer la dimension du noyau de la matrice A ainsi qu'une matrice nulle associée Z lorsque $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$.

17. Cette formule porte le nom de formule *PSB* (pour *Powell's Symmetric Broyden formula*). En pratique, elle donne généralement de moins bons résultats que la formule de BFGS.
18. On observera que ce cas ne peut pas se présenter en optimisation.
19. Ce cas est normalement signe que l'itéré courant est stationnaire.

2. Écrire le système des conditions KKT réduit et en déduire l'unique solution de $(\mathcal{P})$.

Exercice 35. On considère le problème consistant à projeter orthogonalement un point a de $\mathbb{R}^N$ sur le simplexe unité $S = \Big\{x \in \mathbb{R}^N \,|\, x_i \geq 0,\ i = 1, \ldots, N,\ \sum_{i=1}^N x_i \leq 1\Big\}$.
On met sous forme canonique, puis on indexe les $N+1$ contraintes du problème de 1 à $N+1$ dans l'ordre suivant :

$$-x_i \leq 0,\ i = 1, \ldots, N, \text{ et } \sum_{i=1}^N x_i \leq 1,$$

en notant M_i, $i = 1, \ldots, N$, et Λ les multiplicateurs de Lagrange associés, dans le même ordre.

1. Prouver que la $N+1^{\text{ème}}$ contrainte est toujours saturée si $a \geq 0$ et $a \notin S$.
2. On note alors I l'ensemble, contenu dans $\{1, \ldots, N\}$, des indices des contraintes de positivité actives. Montrer que les multiplicateurs associés à la solution du problème *sous les contraintes actives* sont donnés par

$$\Lambda = \frac{1 - \sum_{i \notin I} a_i}{\operatorname{card}(I)},\ M_i = a_i + \Lambda,\ \forall i \in I.$$

3. Résoudre le problème $(\mathcal{P})$ si $N = 5$ et $a = (1, 2.4, 3, 1.7, 0.9)$ en utilisant l'algorithme d'activation de contraintes initialisé avec le point $(1, 0, 0, 0, 0)$.

Exercice 36. On cherche la meilleure approximation d'une fonction $\varphi \in L^2(0,1)$ donnée par la série de Fourier suivante :

$$S_N\,\varphi(t) = \frac{a_0}{\sqrt{2}} + \sum_{k=1}^N a_k\,\cos(k\pi\,t),$$

sous la contrainte que la valeur de $\varphi(t)$ n'excède pas une grandeur donnée δ, dans un sens ou dans l'autre, en certains points t_i, $i = 1, \ldots, p$, et s_j, $j = 1, \ldots, q$, donnés de l'intervalle $[0,1]$, c'est-à-dire [20]

$$(\mathcal{P}) \qquad \min_{\substack{\frac{a_0}{\sqrt{2}} + \sum_{k=1}^N a_k\,\cos(k\pi\,t_i) \leq \varphi(t_i) + \delta,\ i=1,\ldots,p \\ \frac{a_0}{\sqrt{2}} + \sum_{k=1}^N a_k\,\cos(k\pi\,s_j) \geq \varphi(s_j) - \delta,\ j=1,\ldots,q}} \int_0^1 \left[\varphi(t) - \frac{a_0}{\sqrt{2}} - \sum_{k=1}^N a_k\,\cos(k\pi\,t)\right]^2 \mathrm{d}t.$$

1. Expliciter, en fonction des réels a_k, $k = 0, \ldots, N$, le critère de ce problème. On rappelle que, pour tout couple d'entiers naturels (k,l), on a

$$\int_0^1 \cos(k\pi\,t)\,\cos(l\pi\,t)\,\mathrm{d}t = \frac{1}{2}\,\delta_{kl},$$

où δ_{kl} désigne le symbole de Kronecker.

2. Quelle est la nature du problème $(\mathcal{P})$? Justifier l'existence d'une solution.
3. On suppose à présent que $\varphi(t) = \chi_{[0,1/2]}(t)$ et $N = 2$.
   a. Résoudre $(\mathcal{P})$ dans le cas où $p = q = 0$ (pas de contraintes) et représenter l'allure du graphe de la solution $S_2\,\varphi$ trouvée (on indique que $2/\pi \simeq 0.63$) [21].
   b. Résoudre ensuite $(\mathcal{P})$ dans le cas où $p = q = 1$, $t_1 = 0$ et $s_1 = 1$. Donner à nouveau l'allure du graphe de l'unique solution $S_2\,\varphi$.

20. La résolution de ce problème permet de s'affranchir de certaines difficultés liées au phénomène de Gibbs.
21. Le résultat est la série de Fourier tronquée de la fonction paire obtenue en symétrisant le graphe de φ.

Programmation linéaire

Exercice 37.

1. Prouver que le problème linéaire

$$(\mathcal{P}) \qquad \max_{\substack{2\,x_1+x_3+x_4\leq 12\\ x_1+2\,x_2+x_4\leq 12\\ x_i\geq 0,\ i=1,\ldots,4}} (4\,x_1+3\,x_2+x_3+x_4)$$

   admet des solutions.

2. Prouver que l'une au moins des solutions de $(\mathcal{P})$ est un sommet de l'ensemble admissible.
3. Résoudre $(\mathcal{P})$ en listant les sommets de l'ensemble admissible et en comparant les valeurs du critère en ces points.

Exercice 38. On cherche à décider si la distribution de probabilité d'une variable aléatoire pouvant prendre trois valeurs a_i, $i = 1, 2, 3$, est $p = (1/6, 1/3, 1/2)$ ou $q = (1/2, 1/3, 1/6)$, à partir de l'observation d'une seule réalisation de X. On décide de dire que la distribution est p avec une probabilité x_i et q avec une probabilité $1 - x_i$ lorsque la réalisation observée est a_i, $i = 1, 2, 3$. Montrer que le problème consistant à maximiser la probabilité de reconnaître p sous la condition que la probabilité de reconnaître q reste supérieure à 0.95 s'écrit

$$\max_{\substack{3\,x_1+2\,x_2+x_3\leq 0.3\\ x_i\geq 0,\ i=1,2,3}} (x_1+2\,x_2+3\,x_3).$$

Quelle est la règle de décision optimale ?

Exercice 39. Une variable aléatoire discrète X prend trois valeurs possibles -1, 0 ou 1. La distribution de probabilité

$$p = \mathrm{P}(X=-1),\ q = \mathrm{P}(X=0) \text{ et } r = \mathrm{P}(X=1)$$

est inconnue, mais on dispose de l'information :

$$E(X) \in [-0.1, 0.1] \text{ et } E(X^2) \geq 0.9,$$

où $E(X)$ et $E(X^2)$ désignent respectivement les moments d'ordre un et deux de X.

1. Quelles conditions doit vérifier le vecteur (p, q, r) pour définir une distribution de probabilité compatible avec l'information connue *a priori* ? On dira dans la suite qu'un vecteur (p, q, r) est une distribution « admissible » s'il satisfait ces conditions.
2. Vérifier que l'ensemble des distributions admissibles est un polyèdre de $\mathbb{R}^3$. Quels sont ses sommets ?
3. Formuler les problèmes consistant à maximiser et à minimiser $\mathrm{P}(X = 1)$ sachant que $E(X) \in [-0.1, 0.1]$ et $E(X^2) \geq 0.9$. En déduire un encadrement de cette probabilité.

Exercice 40. On considère le problème d'optimisation sans contrainte suivant :

$$(\mathcal{P}) \qquad \min_{(a,b)\in\mathbb{R}^2} \left(|2-a-b| + |4-2\,a-b| + |3-3\,a-b|\right).$$

1. En introduisant trois variables auxiliaires, récrire $(\mathcal{P})$ comme un problème de programmation linéaire.
2. Prouver que l'un des sommets du polyèdre des points admissibles de ce problème est solution. Combien ce polyèdre a-t-il de sommets ? Quelle est la solution de $(\mathcal{P})$?
3. Combien y aurait-t-il de variables et de contraintes si l'on récrivait le problème linéaire précédent sous la forme standard utilisée par l'algorithme du simplexe ?

Exercice 41. On cherche une éventuelle solution $x \geq 0$ du système d'équations $A\,x = b$, où

$$A = \begin{pmatrix} 4 & 1 & 2 & 0 & -1 \\ 0 & 1 & 1 & -1 & 2 \\ 1 & 1 & 1 & -1 & 1 \end{pmatrix} \text{ et } b = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix},$$

en cherchant à résoudre le problème

$$(\mathcal{P}) \qquad \min_{\substack{A\,x+y=b \\ x,y\geq 0}} (y_1 + y_2 + y_3).$$

1. Résoudre $(\mathcal{P})$ par l'algorithme du simplexe initialisé avec $(0,0,1,0,0,1,0,1)$.
2. En conclure que le problème initial n'admet aucune solution.

Exercice 42. On considère le polyèdre Σ de $\mathbb{R}^6$ défini par les inégalités

$$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,\ x_1 + x_4 + x_6 \geq 1,\ x_2 + x_4 + x_5 \geq 1,\ x_3 - x_4 + x_5 + x_6 \geq 1.$$

1. Est-il non vide ? Est-il borné ? Justifier les réponses.
2. Trouver ses éventuels sommets.
3. On cherche à minimiser $f(x) = x_1 + 2\,x_2 + 3\,x_3 - x_4 + x_5 + x_6$ sur Σ.
   a. Écrire les conditions KKT pour le problème correspondant.
   b. Prouver l'existence d'une solution et calculer la valeur du minimum.