Optimisation sous contraintes
Examen

L'examen est prévu pour une durée de 2h30.

La rigueur de la rédaction est un élément d'évaluation important.

Dans tout le problème,

- on note $z \geq 0$ pour signifier qu'un vecteur z a toutes ses composantes positives,
- la norme $\|\cdot\|$ fait référence à la norme euclidienne.

1 Un problème auxiliaire

Pour y et $z \in \mathbb{R}^n$ deux vecteurs donnés, on considère le problème suivant, de variable $x \in \mathbb{R}^n$:

$$\begin{array}{ll} \text{min.} & \frac{1}{2}\|z-x\|^2 \\ \text{s.c.} & x \geq 0, \\ & \langle y, x\rangle = 0. \end{array} \tag{P$_1$}$$

1. Montrer que le problème (P$_1$) est bien posé et admet un unique minimiseur, noté $\varphi_y(z)$.
   Le problème est clairement convexe puisque les contraintes d'inégalité et d'égalité sont affines et la fonction objectif est convexe sur $\mathbb{R}^n$. Elle est même (1-)fortement convexe sur $\mathbb{R}^n$ puisqu'elle s'écrit
   $$x \mapsto \frac{1}{2}\|x\|^2 - \langle z, x\rangle + \frac{1}{2}\|z\|^2,$$
   et que la fonction $x \mapsto -\langle z, x\rangle + \frac{1}{2}\|z\|^2$ est affine (donc convexe).
   La fonction objectif est donc régulière, fortement convexe sur $\mathbb{R}^n$; elle est donc coercive sur $\mathbb{R}^n$. Comme l'ensemble des contraintes est fermé (les fonctions impliquées pour le définir étant continues), le problème est bien posé. Enfin, la forte convexité entraîne la stricte convexité et assure qu'il existe un unique minimiseur.
   Méthode alternative : comme on le fera lors de la partie 2, on peut ici remarquer que ce problème d'optimisation n'est autre que la projection de z sur le convexe fermé $\{x \in \mathbb{R}^n,\ x \geq 0,\ \langle y, x\rangle = 0\}$, qui admet bien un unique minimiseur en la personne du projeté de z sur cet ensemble.
2. Soit $x \in \mathbb{R}^n$ satisfaisant les conditions de KKT pour le problème (P$_1$).
   (a) Montrer qu'il existe un couple $(\lambda, \nu) \in \mathbb{R}^n \times \mathbb{R}$ tel que $\lambda \geq 0$ et
   $$\forall i \in \{1, \ldots, n\}, \quad \lambda_i x_i = 0, \qquad \text{et} \qquad x - z = \lambda - \nu y.$$

   On note f la fonction objectif, $f_1, \ldots, f_n$ les fonctions associées aux contraintes d'inégalité, c'est-à-dire $f_i = x \mapsto -x_i$ pour $i \in \{1, \ldots, n\}$, et $h : x \mapsto \langle y, x\rangle$ la fonction associée à la contrainte d'égalité, on sait qu'il existe $\lambda \in \mathbb{R}^n$, $\nu \in \mathbb{R}$ tels que $\lambda \geq 0$,
   $$\forall i \in \{1, \ldots, n\}, \quad \lambda_i f_i(x) = 0, \qquad \text{et} \qquad \nabla f(x) + \sum_{i=1}^n \lambda_i f_i(x) + \nu \nabla h(x) = 0.$$
   La première condition est exactement celle demandée. Il en va de même pour la seconde puisque pour tout $x \in \mathbb{R}^n$, $\nabla f(x) = x - z$, $\nabla f_i(x) = -e_i$ pour $1 \leq i \leq n$, $\nabla h(x) = y$ et donc
   $$\nabla f(x) + \sum_{i=1}^n \lambda_i f_i(x) + \nu \nabla h(x) = x - z - \sum_{i=1}^n \lambda_i e_i + \nu y = x - z - \lambda + \nu y.$$

On note $I(\nu) = \{i \in \{1,\dots n\},\ \nu y_i - z_i > 0\}$.

(b) En déduire que λ est donné en fonction de ν par

$$\lambda_i = \begin{cases} \nu y_i - z_i & \text{si } i \in I(\nu), \\ 0 & \text{sinon.} \end{cases}$$

Soit $i \in \{1,\dots,n\}$. Supposons pour commencer qu'on ait $i \in I(\nu)$. Alors $x_i - \lambda_i = z_i - \nu y_i < 0$. Ainsi, $x_i < \lambda_i$, mais comme $x_i \geq 0$ on en déduit $\lambda_i > 0$, et alors la condition $x_i \lambda_i = 0$ impose $x_i = 0$, et donc $\lambda_i = x_i + \nu y_i - z_i = \nu y_i - z_i$.

Si maintenant $i \in I(\nu)$, on trouve cette fois $x_i \geq \lambda_i$. Comme $\lambda_i \geq 0$, on a soit $x_i > 0$ auquel cas $\lambda_i = 0$, soit $x_i = 0$ auquel cas $0 \leq \lambda_i \leq x_i = 0$ doit aussi valoir 0.

On note $\lambda(\nu)$ ce vecteur.

(c) Justifier que ν est solution de l'équation

$$\nu = \tfrac{1}{\|y\|^2} \langle y, \lambda(\nu) + z \rangle .$$

Il nous reste à utiliser la condition $\langle y, x \rangle = 0$, qui se traduit en

$$0 = \langle y, z + \lambda(\nu) - \nu y \rangle = \langle y, \lambda(\nu) + z \rangle - \nu \|y\|^2.$$

3. À partir de ce qui précède, proposer une manière de calculer l'unique minimiseur $\varphi_y(z)$ de (P_1) en pratique.

D'après le cours, on sait que tout point qui satisfait les conditions de KKT est un minimiseur du problème puisque le problème est convexe. C'en est donc alors l'unique minimiseur $\varphi_y(z)$. Il suffit ainsi de trouver une solution à l'équation $\nu = \frac{1}{\|y\|^2} \langle y, \lambda(\nu) + z \rangle$, ce qui définit un certain $\lambda(\nu)$ et alors $x = z + \lambda(\nu) - \nu y$ satisfait bien toutes les conditions de KKT, et est donc égal à $\varphi_y(z)$.

Tout revient donc à la résolution d'une équation sur $\mathbb{R}$, qu'on peut envisager de résoudre par dichotomie.

2 Un autre problème auxiliaire

On se donne une matrice $A \in \mathcal{M}_n(\mathbb{R})$ symétrique positive, et un vecteur $b \in \mathbb{R}^n$. On considère le problème

$$\begin{array}{ll} \text{min.} & \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle \\ \text{s.c.} & x \geq 0, \\ & \langle y, x \rangle = 0. \end{array} \tag{P$_2$}$$

1. Donner le gradient et la hessienne de la fonction objectif associée au problème (P_2).

La fonction objectif est de classe C^1 sur $\mathbb{R}^n$ et son gradient est la fonction $x \mapsto Ax - b$. Sa hessienne est la fonction (à valeur matricielle) constante $x \mapsto A$.

2. Proposer un algorithme pour résoudre (P_2), dont on écrira les itérées à partir de la fonction φ_y définie dans la première partie.

   L'ensemble $C_y := \{x \in \mathbb{R}^n,\ x \geq 0,\ \langle y, x\rangle = 0\}$ étant convexe fermé comme on l'a déjà vu, on peut tenter de résoudre ce problème par l'algorithme du gradient projeté. Or la projection d'un vecteur z sur l'ensemble C_y est exactement donné par $\varphi_y(z)$ défini dans la première partie.

   Les itérées s'écrivent alors
$$x_{k+1} = \varphi_y(x_k - \rho(Ax_k - b)),$$
   partant d'un point initial $x_0 \in \mathbb{R}^n$ et pour un pas $\rho > 0$ donné.

3. Donner une hypothèse suffisante portant sur A pour que le problème soit bien posé, admette un unique minimiseur, et que l'algorithme ci-dessus converge vers l'unique minimiseur si le pas est choisi suffisamment petit.

   Pour que le problème soit bien posé et admette un unique minimiseur, il suffit que la fonction objectif soit fortement convexe sur $\mathbb{R}^n$, par un raisonnement analogue à celui fait dans la question 1 de la partie 1. Or cette fonctionnelle quadratique est fortement convexe dès que (en fait, si et seulement si) A est définie positive puisqu'alors $\langle Ax, x\rangle \geq \lambda_m \|x\|^2$ où λ_m est la plus petite valeur propre de A. Comme on suppose que A est déjà positive, une condition suffisante est donc que A soit inversible.

   Par ailleurs, on sait alors sous ces hypothèses que l'algorithme du gradient projeté converge, pour peu que le pas soit suffisamment petit, puisque la fonction à minimiser est fortement convexe, et son gradient est une fonction $\|A\|_2$-lipschitzienne sur $\mathbb{R}^n$.

3 Machine à vecteurs de support

On se donne un jeu de données étiqueté, c'est-à-dire n points x_i de $\mathbb{R}^d$, auxquels sont assignés les étiquettes $y_i \in \{-1, 1\}$ pour $i \in \{1, \dots n\}$. On fera l'hypothèse suivante :

$$\text{les valeurs } -1 \text{ et } 1 \text{ sont chacune représentées au moins une fois au sein du vecteur } y \in \mathbb{R}^n. \tag{H}$$

On se concentre ici sur le cas où il existe un hyperplan qui sépare les données étiquetées $+1$ de celles étiquetées -1 ; quitte à normaliser, cela revient à supposer qu'il existe $w \in \mathbb{R}^d$, $w_0 \in \mathbb{R}$ tels que $\langle w, x_i\rangle - w_0 \geq 1$ si $y_i = 1$ et $\langle w, x_i\rangle - w_0 \leq -1$ si $y_i = -1$. Comme il existera alors en général de très nombreux autres hyperplans qui réaliseront cette séparation, on cherche celui qui maximise la *marge* entre les deux groupes de points, qui n'est autre que la grandeur $\frac{2}{\|w\|}$.

La contrainte ci-dessus s'écrivant encore $y_i(\langle w, x_i\rangle - w_0) \geq 1 \iff 1 - y_i(\langle w, x_i\rangle - w_0) \leq 0$, ces considérations conduisent à résoudre le problème d'optimisation contraint

$$\begin{aligned} \min.\quad & \tfrac{1}{2}\|w\|^2 \\ \text{s.c.}\quad & 1 - y_i(\langle w, x_i\rangle - w_0) \leq 0, \quad i = 1, \dots, n, \end{aligned} \tag{P}$$

qui porte sur la variable $(w, w_0) \in \mathbb{R}^d \times \mathbb{R}$.

1. Montrer que le problème (P) est bien posé.

   On pourra montrer que la fonction objectif $(w, w_0) \mapsto \frac{1}{2}\|w\|^2$ est coercive sur l'ensemble défini par les contraintes, sous l'hypothèse (H).

La fonction à minimiser est continue, l'ensemble des contraintes $S := \{(w, w_0) \in \mathbb{R}^d \times \mathbb{R},\ \forall i \in \{1, \ldots, n\},\ 1 - y_i(\langle w, x_i\rangle - w_0) \leq 0\}$ est fermé. Si on montre que $(w, w_0) \mapsto \frac{1}{2}\|w\|^2$ est coercive sur S, on en déduira que le problème est bien posé.
Soit donc une suite $(w^{(k)}, w_0^{(k)})$ d'élements de S telle que $(\|(w^{(k)}, w_0^{(k)})\|)$ tend vers $+\infty$. Si par l'absurde la fonction objectif ne tend pas vers $+\infty$ le long de cette suite, c'est qu'on peut extraire une sous-suite (via une extractrice φ) telle que $(\|w^{(\varphi(k))}\|)$ est bornée. Comme $\|(w^{(k)}, w_0^{(k)})\| = \|w^{(k)}\| + |w_0^{(k)}|$ pour tout $k \in \mathbb{N}$, c'est donc que $(|w_0^{(\varphi(k))}|)$ tend vers $+\infty$. Il est donc possible d'extraire une nouvelle sous-suite de la première (notons ψ l'extractrice obtenue après ces deux extractions) telle que soit $(w_0^{(\psi(k))})$ tend vers $+\infty$, soit $(w_0^{(\psi(k))})$ tend vers $-\infty$.
On examine les inégalités

$$1 - y_i(\langle w^{(\psi(k))}, x_i\rangle - w_0^{(\psi(k))}) \leq 0,$$

valables pour tout $i \in \{1, \ldots, n\}$, $k \in \mathbb{N}$ puisque les suites sont dans S. Si $(w_0^{(\psi(k))})$ tend vers $+\infty$, on choisit i tel que $y_i = 1$. Le terme en produit scalaire est borné, alors que le dernier terme tend vers $+\infty$, ce qui contredit la majoration. Si $(w_0^{(\psi(k))})$ tend vers $-\infty$, on choisit i tel que $y_i = -1$, aboutissant là aussi à une contradiction.
Finalement, la fonctionnelle objectif est bien coercive sur S.

2. On s'intéresse au problème dual au problème (P) ; **on admettra qu'il y a dualité forte et que l'optimum du problème dual est atteint s'il est fini.**
   (a) Montrer que le Lagrangien $L : (w, w_0, \lambda) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^n$ associé à (P) est donné par

$$\forall (w, w_0, \lambda) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^n, \quad L(w, w_0, \lambda) = \tfrac{1}{2}\|w\|^2 - \left\langle w, \sum_{i=1}^n y_i \lambda_i x_i \right\rangle + \langle \lambda, y\rangle\, w_0 + \sum_{i=1}^n \lambda_i$$

   Par définition, on trouve

$$\begin{aligned}\forall (w, w_0, \lambda) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^n, \quad L(w, w_0, \lambda) &= \tfrac{1}{2}\|w\|_2^2 + \sum_{i=1}^n \lambda_i (1 - y_i(\langle w, x_i\rangle - w_0)) \\ &= \tfrac{1}{2}\|w\|^2 - \left\langle w, \sum_{i=1}^n y_i \lambda_i x_i \right\rangle + \langle \lambda, y\rangle\, w_0 + \sum_{i=1}^n \lambda_i\end{aligned}$$

   (b) Montrer que pour $\lambda \in \mathbb{R}^n$ fixé, $\inf_{(w,w_0)\in\mathbb{R}^d\times\mathbb{R}} L(w, w_0, \lambda) > -\infty$ si et seulement si $\langle y, \lambda\rangle = 0$, et qu'alors $(w(\lambda), w_0(\lambda))$ est un minimiseur de $(w, w_0) \mapsto L(w, w_0, \lambda)$ sur $\mathbb{R}^d \times \mathbb{R}$ si et seulement si $w(\lambda) = \sum_{j=1}^n y_j \lambda_j x_j$.
   Pour $\lambda \in \mathbb{R}^n$ fixé, la fonction $(w, w_0) \mapsto L(w, w_0, \lambda)$ est la somme de deux fonctions, l'une qui ne dépend que de w donnée par $w \mapsto \frac{1}{2}\|w\|^2 - \langle w, \sum_{i=1}^n y_i \lambda_i x_i\rangle$, l'autre que de w_0 donnée par $w_0 \mapsto \langle \lambda, y\rangle\, w_0 + \sum_{i=1}^n \lambda_i$. Pour que l'infimum ne soit pas $-\infty$, il faut et il suffit donc que l'infimum de chacune de ces deux fonctions (en $w \in \mathbb{R}^d$ et $w_0 \in \mathbb{R}$, respectivement) ne soient pas $-\infty$. Or la deuxième fonction est affine, et est donc minorée (et alors constante) si et seulement si $\langle \lambda, y\rangle = 0$. La première fonction est, elle, clairement fortement convexe donc sa minimisation conduit à un problème bien posé.
   Toujours du fait de cette remarque, lorsque $\langle y, \lambda\rangle = 0$, $(w(\lambda), w_0(\lambda))$ est alors un minimiseur de $(w, w_0) \mapsto L(w, w_0, \lambda)$ si et seulement si $w(\lambda)$ est un minimiseur de $w \mapsto \frac{1}{2}\|w\|^2 - \langle w, \sum_{i=1}^n y_i \lambda_i x_i\rangle$. L'unique minimiseur de cette fonction est donné par l'unique point qui annule son gradient donné par $w \mapsto w - \sum_{i=1}^n y_i \lambda_i x_i$. C'est donc bien que $w(\lambda) = \sum_{j=1}^n y_j \lambda_j x_j$.

(c) En déduire que le problème dual peut s'écrire

$$\begin{aligned} \text{max.} \quad & -\tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \langle x_i, x_j\rangle \lambda_i\lambda_j + \sum_{i=1}^{n}\lambda_i \\ \text{s.c.} \quad & \lambda \geq 0 \\ & \langle y, \lambda\rangle = 0. \end{aligned} \tag{D}$$

Notons g la fonction duale du problème. On a vu que $g(\lambda) = -\infty$ si et seulement si $\langle y, \lambda\rangle = 0$. Cette contrainte peut être rendue explicite via la formulation

$$\begin{aligned} \text{max.} \quad & g(\lambda) \\ \text{s.c.} \quad & \lambda \geq 0 \\ & \langle y, \lambda\rangle = 0. \end{aligned}$$

Dans le cas où $\langle y, \lambda\rangle = 0$, et pour $(w(\lambda), w_0(\lambda))$ un minimiseur de $(w, w_0) \mapsto L(w, w_0, \lambda)$, la question précédente permet en outre d'écrire

$$\begin{aligned} g(\lambda) = L(w(\lambda), w_0, \lambda) &= \tfrac{1}{2}\|w(\lambda)\|^2 - \left\langle w(\lambda), \sum_{i=1}^{n} y_i\lambda_i x_i \right\rangle + \sum_{i=1}^{n}\lambda_i \\ &= \tfrac{1}{2}\|w(\lambda)\|^2 - \|w(\lambda)\|^2 + \sum_{i=1}^{n}\lambda_i = -\tfrac{1}{2}\|w(\lambda)\|^2 + \sum_{i=1}^{n}\lambda_i \\ &= -\tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \langle x_i, x_j\rangle \lambda_i\lambda_j + \sum_{i=1}^{n}\lambda_i. \end{aligned}$$

(d) Soit $\lambda^\star \in \mathbb{R}^n$ une variable duale optimale et $(w^\star, w_0^\star) \in \mathbb{R}^d \times \mathbb{R}$ une variable primale optimale. Montrer les relations

- ⋄ $w^\star = \sum_{j=1}^{n} y_j \lambda_j^\star x_j$,
- ⋄ pour $i \in \{1, \dots, n\}$, si $\lambda_i^\star > 0$, alors $y_i(\langle w^\star, x_i\rangle - w_0^\star) = 1$.

Comme il y a dualité forte, la donnée de variables primale et duale optimales permet d'affirmer que $(w^\star, w_0^\star, \lambda^\star)$ est un point-selle du Lagrangien.

D'une part, $(w^\star, w_0^\star)$ est un minimiseur de $(w, w_0) \mapsto L(w, w_0, \lambda^\star)$ sur $\mathbb{R}^d \times \mathbb{R}$, ce qui impose par les questions précédentes $\langle y, \lambda^\star\rangle = 0$, et $w^\star = w(\lambda^\star) = \sum_{j=1}^{n} y_j \lambda_j^\star x_j$.

D'autre part, $\lambda^\star$ est un maximiseur de $\lambda \mapsto L(w^\star, w_0^\star, \lambda)$ sur $\mathbb{R}_+^n$, c'est-à-dire qu'il s'agit d'un maximiseur de $\lambda \mapsto \sum_{i=1}^{n} \lambda_i(1 - y_i(\langle w^\star, x_i\rangle - w_0^\star))$. Tous les termes de cette somme sont négatifs et cette somme admet pour maximum 0, atteinte donc si tous les termes de la somme sont nuls. En particulier, si $\lambda_i^\star > 0$, alors on doit avoir $(\langle w^\star, x_i\rangle - w_0^\star) = 1$.

Méthode alternative : puisqu'il y a dualité forte, on peut écrire les conditions de KKT satisfaites par $(w^\star, w_0^\star)$ avec $\lambda^\star$ comme multiplicateur de Lagrange associé.

3. On discute désormais d'une méthode algorithmique globale pour résoudre (P).

(a) Peut-on utiliser l'algorithme d'Uzawa pour résoudre (P) ?

Comme le problème de minimisation du Lagrangien n'est pas bien posé pour n'importe quelle valeur de $\lambda \in \mathbb{R}_+^n$, l'algorithme d'Uzawa n'est pas défini pour ce problème.

(b) Justifier que le problème dual (D) se ramène à un problème de la forme (P_2), en précisant la matrice $A \in \mathcal{M}_n(\mathbb{R})$ et le vecteur $b \in \mathbb{R}^n$ impliqués.

On constate en effet que (D) est exactement de la forme (P_2) après avoir considéré le problème de minimisation de l'opposée de la fonction obtenue, qui correspond à la matrice A définie par $a_{i,j} = y_i y_j \langle x_i, x_j \rangle$, et le vecteur b défini par par $b_i = 1$ pour tous $i, j \in \{1, \dots, d\}$. La matrice A, symétrique, est bien positive puisque pour $u \in \mathbb{R}^n$, $\langle Au, u \rangle = \sum_{i=1}^n \sum_{j=1}^n y_i y_j \langle x_i, x_j \rangle u_i u_j = \| \sum_{i=1}^n y_i u_i x_i \|^2 \geq 0$.

(c) Conclure alors à un algorithme de résolution du problème dual (D), et expliquer comment en déduire une estimation d'un minimiseur de (P).

D'après tout ce qui précède, notant A et b la matrice et le vecteur définis par la question précédente, on utilise l'algorithme du gradient projeté (où la projection est obtenue par évaluation de la fonction φ_y définie dans la première partie). Sous réserve de convergence (par exemple si les données sont telles que A est définie positive) vers une variable duale optimale, on en déduit une estimation $\tilde{\lambda}$ d'une telle variable.

En résulte une estimation d'une variable primale optimale $(w^\star, w_0^\star)$. En effet, on pose suivant la question 2.(d) $\tilde{w} := \sum_{j=1}^n y_j \tilde{\lambda}_j x_j$. Puis en considérant un indice i_0 tel que $\tilde{\lambda}_{i_0} > 0$ et toujours motivé par la question 2.(d), on pose $\tilde{w}_0 := \langle \tilde{w}, x_{i_0} \rangle + y_{i_0}$.