

1 Optimisation sans contraintes

Exercice 1 Calcul différentiel. (★)

1. Calculez les dérivées partielles par rapport à x et à y de la fonction $f : x, y \mapsto x^2y$ de \mathbf{R}^2 vers \mathbf{R} .
2. Utilisez la règle donnant le Jacobien d'une fonction composée pour calculer le gradient de la fonction $x \mapsto \|x\|_2$ de \mathbf{R}^n vers \mathbf{R} .
3. Calculez la Hessienne de la fonction $f : x, y \mapsto x^2 + y^2$ de \mathbf{R}^2 vers \mathbf{R} .
4. Soit M une matrice à coefficients réels de taille n par n . Calculez la Hessienne de la fonction $g_M : x \mapsto x^T Mx$ de \mathbf{R}^n vers \mathbf{R} .

Exercice 2 Existence et unicité des extrema. (★)

1. Considérons $f : x \mapsto \sin(x)$ de $[0, \pi]$ vers \mathbf{R} . On admet que comme f est définie sur un intervalle fermé borné et continue sur cet intervalle, elle admet un minimum global (cf. théorème de Weierstrass en analyse réelle https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me_des_valeurs_extr%C3%AAmes). Trouver le minimum global de f en utilisant une condition nécessaire d'existence de minimum global.
2. Montrez que la fonction $f : x, y \mapsto x^2 + y^2$ de \mathbf{R}^2 vers \mathbf{R} admet un unique minimum global et identifiez ce minimum.
3. Prouvez que la fonction $f : x_1, x_2 \mapsto x_1^2 - 36x_1 - 36x_2 + x_2^2 + 150$ de \mathbf{R}^2 vers \mathbf{R} admet un minimum global.

Exercice 3 Déterminez si $f : x, y \mapsto xy + x^2 + 1 + 2y$ est convexe.

Exercice 4 Preuve de l'existence de la décomposition en valeurs singulières. (★★)

1. Prouver que toute matrice semi-orthogonale de taille n par r , avec $r \leq n$, peut être complétée en une matrice orthogonale de taille n par n .

2. Soit A une matrice réelle de taille m par n . Montrer qu'il est possible de trouver une matrice colonne x de taille n et une matrice colonne y de taille m , telles que : $Ax = \sigma y$, $\|x\|_2 = 1$, $\|y\|_2 = 1$ et $\sigma = \|A\|_2$. Indice : utiliser l'une des deux définitions équivalentes de $\|A\|_2$.
3. En utilisant le résultat de la première question, x et y , montrer qu'il existe une matrice orthogonale U de taille m par m , une matrice orthogonale V de taille n par n , une matrice colonne w de taille $n - 1$ et une matrice B de taille $m - 1$ par $n - 1$, telles que :

$$A_1 := U^T A V = \begin{pmatrix} \sigma & w^T \\ 0 & B \end{pmatrix}.$$

4. Montrer que $\|A_1\|_2^2 \geq \sigma^2 + w^T w$. Indice : multiplier A_1 par $(\sigma w^T)^T$.
5. Montrer que $\|A_1\|_2^2 = \|A\|_2^2$ et en déduire la valeur de w .
6. Utiliser un raisonnement par récurrence pour compléter la preuve de l'existence de la décomposition en valeurs singulières.

2 Optimisation sous contraintes

Point de cours. Conditions de Karush-Kuhn-Tucker pour des fonctions \mathcal{C}^1 .

Soit f , $(e_i)_{i=1}^{n_E}$ et $(c_i)_{i=1}^{n_I}$ des fonctions de \mathbf{R}^d vers \mathbf{R} de classe \mathcal{C}^1 .

Comme nous l'avons vu en cours, on dit que le point $x^* \in \mathbf{R}^d$ vérifie les conditions de Karush-Kuhn-Tucker (KKT) pour le problème :

$$\min_{x \in \mathbf{R}^d} f(x) \text{ tel que } \begin{cases} e_i(x) = 0, & i \in \mathcal{E} \\ c_i(x) \leq 0, & i \in \mathcal{I} \end{cases}$$

si et seulement si il existe $\lambda_E^* \in \mathbf{R}^{n_E}$ et $\lambda_I^* \in \mathbf{R}^{n_I}$, tels que :

1. $\nabla_x \mathcal{L}(x^*, \lambda_E^*, \lambda_I^*) = 0$, où $\mathcal{L}(x, \lambda_E, \lambda_I) = f(x) + \sum_{i=1}^{n_E} \lambda_{E,i} e_i(x) + \sum_{i=1}^{n_I} \lambda_{I,i} c_i(x)$
2. $e_i(x^*) = 0$ pour tout i dans $\{1, 2, \dots, n_E\}$
3. $c_i(x^*) \leq 0$ pour tout i dans $\{1, 2, \dots, n_I\}$
4. $\lambda_{I,i} \geq 0$ pour tout i dans $\{1, 2, \dots, n_I\}$
5. $\lambda_{I,i} c_i(x^*) = 0$ pour tout i dans $\{1, 2, \dots, n_I\}$

Les trois théorèmes suivants donnent des conditions de régularité sur les fonctions du problème—appelées *qualifications des contraintes*—qui garantissent que toute solution locale du

problème considéré doit vérifier les conditions KKT (c'est à dire que sous ces conditions de régularité les conditions KKT sont des conditions nécessaires de solution locale). Notez qu'on peut démontrer la validité d'autres qualification des contraintes que les trois présentées ici et qu'il est possible de généraliser ces théorèmes au delà du cas des fonctions de classe \mathcal{C}^1 .

Theorem 1 *Conditions nécessaires de solution locale pour des contraintes linéaires. Si les fonctions e_i et c_j sont affines pour $i = 1 \dots n_E$ et $j = 1 \dots n_I$ alors les conditions KKT sont des conditions nécessaires de solution locale pour le problème considéré.*

Theorem 2 *Conditions nécessaires de solution locale pour des contraintes linéairement indépendantes.*

*Si au point $x \in \mathbf{R}^d$ la matrice formée par les gradients des contraintes d'égalité et des contraintes d'inégalité **actives** est de rang plein (c'est à dire que les n_E vecteurs gradients pour les contraintes d'égalité—qui doivent toujours toutes être actives—et les vecteurs gradients pour les contraintes d'inégalité actives en x —c'est à dire telles que $c_i(x) = 0$ —sont linéairement indépendants), alors les conditions KKT sont des conditions nécessaires de solution locale en x pour le problème considéré.*

Theorem 3 *Conditions nécessaires de solution locale pour un problème convexe.*

Si f est convexe, c_i est concave pour tout i in $1 \dots n_I$ et e_j est affine pour tout j in $1 \dots n_E$ et si la condition de Slater est vérifiée, c'est à dire s'il existe $x \in \mathbf{R}^d$ tel que $e_j(x) = 0$ pour tout j in $1 \dots n_E$ et $c_i(x) < 0$ pour tout i in $1 \dots n_I$, alors les conditions KKT sont des conditions nécessaires de solution locale pour le problème considéré.

Exercice 5 Application des conditions KKT. (★)

On considère le problème

$$\min_{x,y \in \mathbf{R}^2} x^2 + y^2 \text{ tel que } \begin{cases} x + y \geq 1 \\ y \leq 2 \\ y^2 \geq x \end{cases}$$

1. Donnez les conditions KKT pour ce problème.
2. Trouvez tous les points solutions des conditions de KKT pour ce problème.
3. Donnez en le justifiant toutes les solutions globales de ce problème.
4. Faites un graphe pour vérifier la plausibilité de vos réponses.

Exercice 6 Boîte de surface maximale à diagonale fixée. (★)

Trouvez le parallélépipède rectangle (pavé droit) dont la surface (somme des aires des six faces) est maximale parmi tous les parallélépipèdes rectangles de diagonale $x_1^2 + x_2^2 + x_3^2 = L$, où L est un nombre réel strictement positif fixé et x_1, x_2, x_3 sont les longueurs des côtés du parallélépipède rectangle.

3 Analyse convexe et dualité

Exercice 7 (★) Dualité et contraintes séparables.

On considère un producteur de glace qui souhaite produire au moins 100kg de glace par jour. Ce producteur a trois employés : le premier produit 5kg de glace par heure, le second 10kg par heure et le troisième 1kg par heure. Une journée de travail dure au maximum 8h. Le producteur souhaite minimiser les coûts de production, sachant que l'employé i est payé à la fin de la journée x_i^2 centimes d'euros où x_i est la quantité de glace produite dans la journée par l'employé i .

On modélise ce problème comme suit :

Minimiser la fonction :

$$f : x_1, x_2, x_3 = x_1^2 + x_2^2 + x_3^2$$

sur l'ensemble $X = [0, 40] \times [0, 80] \times [0, 8] \subset \mathbf{R}^3$, sous la contrainte :

$$x_1 + x_2 + x_3 \geq 100$$

1. Justifiez que le problème considéré possède une solution globale.
2. Montrez que la fonction duale du problème considéré peut s'écrire sous la forme :

$$q(\lambda) = \inf_{x_1 \in [0, 40]} f_1(x_1, \lambda) + \inf_{x_2 \in [0, 80]} f_2(x_2, \lambda) + \inf_{x_3 \in [0, 8]} f_3(x_3, \lambda) + 100\lambda$$

pour des fonctions f_1, f_2 et f_3 que vous explicitez.

3. On note $X_1 = [0, 40]$, $X_2 = [0, 80]$, et $X_3 = [0, 8]$. Pour $i \in \{1, 2, 3\}$, donnez $x_i^* = \arg \min_{x_i \in X_i} f_i(x_i, \lambda)$ en fonction de λ , vous pouvez distinguer plusieurs cas de figure à chaque fois.
4. Déduisez-en des formules explicites pour la fonction duale sur chacun des intervalles suivants pour $\lambda : [0, 16], [16, 80], [80, 160], [160, +\infty[$.

5. En utilisant le résultat de la question précédente, résolvez le problème dual.
6. Justifiez que la dualité forte s'applique à ce problème et utilisez les conditions d'optimalité vues en cours pour en déduire une solution du problème primal.
7. Voyez-vous une méthode plus simple qui aurait pu vous permettre de trouver cette solution ?
8. Répondez à nouveau aux questions 1 à 7 en modifiant le problème comme suit : le premier employé est maintenant payé $4x_1^2$ centimes d'euros à la fin de la journée, le second employé $2x_2^2$ centime d'euros et le troisième $10x_3^2$ centime d'euros.

Exercice 8 Écart dual pour le problème du sac à dos. (★★)

On considère n objets avec des poids associés w_1, w_2, \dots, w_n , et des valeurs associées v_1, v_2, \dots, v_n . On veut sélectionner un sous-ensemble de ces objets tel que la somme des poids associés aux éléments de ce sous-ensemble ne dépasse pas un réel $A > 0$ donné et tel que la somme des valeurs associées aux éléments de ce sous-ensemble soit maximisée.

1. Montrer que le problème peut être formulé comme suit :

$$\text{minimiser } f(x) = - \sum_{i=1}^n v_i x_i \quad (\text{Problème du sac à dos})$$

$$\text{avec } \sum_{i=1}^n w_i x_i \leq A$$

et pour tout $i \in \{1, 2, \dots, n\}$, $x_i \in \{0, 1\}$.

2. Trouver graphiquement la solution du problème dual et représenter graphiquement l'écart de dualité pour $n = 5$, $(w_1, w_2, w_3, w_4, w_5) = (1, 2, 3, 4, 5)$ et $(v_1, v_2, v_3, v_4, v_5) = (3, 1, 6, 2, 5)$.
3. Calculer explicitement la fonction duale dans le cas général.
4. Trouver la solution du problème dual. (Indice : on peut faire l'hypothèse, sans perte de généralité, que $\frac{v_1}{w_1} \leq \frac{v_2}{w_2} \leq \dots \leq \frac{v_n}{w_n}$ et étudier la valeur de la fonction duale en fonction de la position de son argument par rapport aux v_i/w_i .)

5. On considère maintenant le problème relaxé :

$$\text{minimiser } f(x) = - \sum_{i=1}^n v_i x_i \quad (\text{Problème du sac à dos relaxé})$$

$$\text{avec } \sum_{i=1}^n w_i x_i \leq A$$

et pour tout $i \in \{1, 2, \dots, n\}$, $x_i \in [0, 1]$.

Justifier qu'il n'y a pas d'écart de dualité pour ce problème.

6. Montrer que le problème dual du problème relaxé atteint le même maximum que le problème dual du problème original.
7. Utiliser les conditions d'optimalité primale-duale vues en cours pour obtenir la solution du problème primal relaxé.
8. Montrer que cette solution est faisable pour le problème primal original et en déduire que :

$$q^* \leq f^* \leq q^* + \max_{1 \leq i \leq n} v_i,$$

où f^* est la solution du problème primal original et q^* est la solution du problème dual original.