

Optimisation pour l'apprentissage automatique

M2 IASD Apprentissage

Projet 2025/2026



Instructions

- Version courante : 28 mai 2026.
- La version la plus récente du projet (mise à jour en cas de changement) est disponible à l'adresse <https://www.lamsade.dauphine.fr/~croyer/ensdocs/OAA/ProjOAA.zip>.
- Le travail décrit ci-dessous peut être réalisé en binôme ou individuellement.
- Il n'y a pas de contraintes sur l'implémentation, mais celle-ci fera partie intégrante du rendu. Un point de départ (ou plutôt de fin) est disponible dans le notebook lié au sujet. Les première et dernière parties peuvent aussi être traitées de manière similaire aux notebooks proposés en séance.
- Les réponses aux différentes questions ainsi que l'implémentation utilisée devront être envoyées au format ZIP à clement.royer@lamsade.dauphine.fr. Le nom de l'archive devra comporter le(s) nom(s) de chaque étudiant(e) impliqué(e) dans le projet.
- La date limite de rendu est actuellement fixée au **28 juin 2026 AOE** (Anywhere On Earth).

Notations

- Les dimensions des vecteurs ou matrices seront toujours supposées supérieures ou égales à 1.
- La notation $\|\cdot\|$ désignera la norme euclidienne d'un vecteur, càd $\|v\| = \sum_{i=1}^m [v]_i^2$ pour tout $v \in \mathbb{R}^m$.
- Pour deux matrices $A, B \in \mathbb{R}^{m_1 \times m_2}$, on notera $\langle A, B \rangle = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} A_{ij} B_{ij}$.
- Le vecteur de \mathbb{R}^d dont toutes les composantes sont égales à 1 sera noté $\mathbf{1}_d$.
- On notera $\text{proj}_{\geq 0}[w]$ la projection du vecteur w sur l'orthant positif.

Introduction : Transfert de couleurs

Ce court projet porte sur le contenu de la séance numéro 7 sur les méthodes primales-duales. Le but du projet est de tester le potentiel de ces méthodes sur une tâche de traitement d'image appelée *transfert de couleurs*. Il s'agit d'appliquer la palette de couleurs d'une image à une autre, comme le montre la figure 1.



Figure 1: Transfert de couleur entre deux images réelles (et libres de droits).

Méthodes primales-duales de base

On commence dans cette partie par considérer des programmes linéaires sous la forme classique

$$\underset{w \in \mathbb{R}^d}{\text{minimiser}} \quad f(w) = c^T w \quad \text{sous contraintes} \quad Aw = b, \quad w \geq 0 \quad (1)$$

avec $A \in \mathbb{R}^{r \times d}$, $b \in \mathbb{R}^r$ et $c \in \mathbb{R}^d$.

Comme vu en cours, une formulation primale-duale de ce problème est

$$\underset{\substack{w \in \mathbb{R}^d \\ w \geq 0}}{\text{minimiser}} \quad \max_{z \in \mathbb{R}^r} \mathcal{L}(w, z) := c^T w - z^T (Aw - b), \quad (2)$$

où l'idée est d'optimiser à la fois sur w et z via des techniques basées sur le gradient.

Question 1 Implémenter la descente/montée de gradient appliquée au problème (2), dont l'itération est donnée par

$$\begin{cases} w_{k+1} &= \text{proj}_{\geq 0} [w_k - \alpha_k \nabla_w \mathcal{L}(w_k, z_k)] = \text{proj}_{\geq 0} [w_k - \alpha_k (c - A^T z_k)] \\ z_{k+1} &= z_k + \beta_k \nabla_z \mathcal{L}(w_k, z_k) = z_k - \beta_k (Aw_k - b). \end{cases} \quad (3)$$

Appliquer la méthode au problème $\min_{w \geq 0} \max_{z \in \mathbb{R}} (w-3)z$ avec $(w_0, z_0) = (2, 2)$ et $\alpha_k = \beta_k = 0.2$. Que remarque-t-on ?

Question 2 Comparer la méthode de la question 1 avec la variante

$$\alpha_{2k} = -\beta_{2k} = -\alpha_{2k+1} = \beta_{2k+1} = 0.2$$

qui a été récemment proposée dans la littérature. En quoi cette proposition est-elle surprenante ?

Question 3 Implémenter la méthode primale-duale de gradient hybride (PDHG) appliquée au problème (2), dont l'itération est donnée par

$$\begin{cases} w_{k+1} &= \text{proj}_{\geq 0} [w_k - \alpha_k (c - A^T z_k)] \\ z_{k+1} &= z_k - \beta_k (A(2w_{k+1} - w_k) - b). \end{cases} \quad (4)$$

Comparer la méthode PDHG aux variantes de la question (1) sur le même problème.

Transport optimal et transfert de couleurs

Étant données deux mesures, c'est-à-dire (pour simplifier) deux vecteurs $\mathbf{a} \in \mathbb{R}^d$ et $\mathbf{b} \in \mathbb{R}^d$ à coefficients positifs ou nuls tels que $\mathbf{a}^T \mathbf{1}_d = \mathbf{b}^T \mathbf{1}_d = 1$, le transport optimal de \mathbf{a} vers \mathbf{b} s'écrit

$$\underset{\mathbf{P} \in \mathbb{R}^{d \times d}}{\text{minimiser}} \langle \mathbf{C}, \mathbf{P} \rangle \quad \text{sous contraintes} \quad \mathbf{P} \mathbf{1}_d = \mathbf{a}, \quad \mathbf{P}^T \mathbf{1}_d = \mathbf{b}, \quad \mathbf{P} \geq \mathbf{0}, \quad (5)$$

où \mathbf{C} est une matrice de coûts représentant le coût de transport des a_i vers les b_j .

Question 4 On considère le problème de transport optimal (5) avec

$$\mathbf{a} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 5 \\ 1 & 2 \end{bmatrix}. \quad (6)$$

Comparer la performance des méthodes GDA et PDHG sur ce problème.

On formule maintenant le problème de transfert de couleur comme un programme linéaire de transport optimal. Étant données deux images I_s et I_t dont les pixels sont normalisés, on sous-échantillonne un nombre d de pixels dans chacune des images, que l'on notera X_s et X_t . On pose alors $\mathbf{a} = \mathbf{b} = \mathbf{1}_d$ et on calcule la matrice de coût $\mathbf{C}_{ij} = \| [X_s]_i - [X_t]_j \|$.

Question 5 Adapter le code fourni dans le projet pour pouvoir appliquer les méthodes GDA et PDHG développées plus haut au problème de transfert de couleur.

Question 6 Comparer les variantes obtenues en question 5 à la méthode utilisée dans le code fourni.

Extension

L'algorithme de Sinkhorn (utilisé en comparaison à la question 6) est une méthode de gradient basée sur une version régularisée du problème de départ. Dans cet esprit, on considère une variante relâchée du transport optimal (dite déséquilibrée) où on remplace les contraintes par des régularisations, à savoir

$$\underset{\substack{\mathbf{P} \in \mathbb{R}^{d \times d} \\ \mathbf{P} \geq \mathbf{0}}}{\text{minimiser}} \langle \mathbf{C}, \mathbf{P} \rangle + \frac{\lambda}{2} \|\mathbf{P} \mathbf{1}_d - \mathbf{a}\|^2 + \frac{\lambda}{2} \|\mathbf{P}^T \mathbf{1}_d - \mathbf{b}\|^2, \quad (7)$$

avec $\lambda > 0$.

Question 7 En se basant sur les méthodes vues en cours, proposer deux méthodes de gradient pour traiter le problème (7). Appliquer ces méthodes aux problèmes jouets des questions 1 et 6.

Question 8 Comparer les méthodes développées en question 7 sur le problème de transfert de couleur.