Projet Optimisation

LIAO Puwei
ZHANG Minghe

January 24, 2021

1 Introduction

Dans ce projet, notre but est de modéliser le transport optimal et d'essayer d'estimer les solutions optimaux pour par exemple le problème de transport entre les boulangeries et les cafés et ainsi les clients.a production et la consommation de pains au chocolat sont décrites par μ et ν,μ_i donne la quantité de pains au chocolat produite par la boulangerie i,ν_j pour le cafe j,notre contraintes de ce probleme est donc:

$$\sum_i \mu_i = \sum_j \nu_j = 1$$

On voudrait trouver une moyenne de sorte que le coût total de transport soit minimum avec un coût C_{ij}, on l'appelle le problème de Monge-Kantorovich, ce qui s'ecrit sous la forme:

$$\min \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} C_{ij}\gamma_{ij} \mid \gamma \in \Pi(\mu, \nu) \right\}$$

2 EXO1

2.1 1.1

On voudrait écrire les contraintes suivantes sous la forme Ax=b

$$\sum_{j=1}^{N} \gamma_{ij} = \mu_i, \text{ et } \sum_{i=1}^{N} \gamma_{ij} = \nu_j$$

Ici la matrice A de taille 2N * (N*N) est décrit de la forme suivante

$$\begin{pmatrix}
1 & . & . & 1 & & & & & & & & & & \\
 & & & & 1 & . & . & 1 & & & & & & \\
 & & & & & & & & . & . & & & & \\
 & & & & & & & & & & 1 & . & . & 1 \\
1 & & & & 1 & & & & & & 1 & & & \\
 & . & & & & . & & & & & & . & & \\
 & & . & & & & . & & . & . & & & . & \\
 & & & 1 & & & & 1 & & & & & & 1
\end{pmatrix}$$

D'ici on peut bien écrire ce problème comme un problème d'optimisation sous contraintes d'égalités

$$\min\{< \boldsymbol{c}, x >\mid A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$$

2.2 1.4

Dans cette question, on utilise lemme de Forkas pour appliquer le probleme dual de MK, qui dit que le probleme précédente est équivalent à problème suivante

$$max\{< y, b > |A^T y \leq c\} \;\; avec \;\; y = \begin{pmatrix} u_1 \\ \vdots \\ u_N \\ v_1 \\ \vdots \\ v_N \end{pmatrix}$$

et on a donc

$$< y, b >= \sum_i u_i \mu_i + \sum_j v_j \nu_j$$

et ainsi $A^T y \leq c$ c'est la meme chose de dire que $u_i + v_j \leq C_{ij}$

donc on peut conclure que le probleme duale est donnée par:

$$\max \left\{ \sum_i u_i \mu_i + \sum_j v_j \nu_j \mid u, v \in \mathbb{R}^N u_i + v_j \leq C_{ij} \forall (i, j) \in I \times J \right\}$$

3 EXO2

3.1 2.1.1

Posons d'abord la fonction log sum exp pour pouvoir avancer dans cette question d'optimalité

$$Soit\ t \in [0, 1], x \in \mathbb{R}^N, y \in \mathbb{R}^N, f(tx + (1 - t)y) = \log\left(\sum_{i=1}^{N} e^{tx_i} e^{(1-t)y_i}\right)$$

On voudrais d'abord étudier la convexité de cette fonction Notons $u_i = e^{x_i} \quad v_i = e^{y_i}$ D'après l'inégalité holder:

$$\sum_{i=1}^{n} x_i y_i \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$$

On obtient alors donc:

$$\log\left(\sum_{i=1}^{n} u_i^t v_i^{(1-t)}\right) \leqslant \underbrace{\log\left[\left(\sum_{i=1}^{n} u_i^{t \frac{1}{t}}\right)^{t} \cdot \left(\sum_{i=1}^{n} v_i^{1-t \cdot \frac{1}{1-t}}\right)^{1-t}\right)}_{t \log\left(\sum_{i=1}^{n} u_i\right) + (1-t) \log\left(\sum_{i=1}^{n} v_i\right).}$$

D'ou on a $f(tx + (1 - t)y) \leqslant tf(x) + (1 - t)f(y)$,ce qui prouve donc la convexité de la fonction f.

3.2 2.1.2

On sait que

$$max\{y_1....y_N\} \leq y_1 + ... + y_N \leq Nmax\{y_1....y_N\}$$

Ici, on met $y_i = e^{x_i}$,on obtient l'inégalité suivante :

$$\max\left(e^{x_i}\right) \leqslant \sum_{i=1}^{N} e^{x_i} \leqslant N \max\left(e^{x_i}\right)$$

en raison de monotonalité croissance de fonction exponentiel, l'inégalité précédente peut se transformer à l'inégalité suivante:

$$\exp\left(\max x_i\right) \leqslant \sum_{i=1}^{N} e^{x_i} \leqslant N \exp\left(\max x_i\right).$$

En mettant la fonction log:

$$\max x_i \leqslant \log(\sum_{i=1}^{N} \exp(x_i)) \leqslant \max x_i + \log N$$

3.3 2.1.3

En mettant $x_i = \frac{x_i}{\varepsilon}$ on peut avoir:

$$\frac{\max x_i}{\varepsilon} \leqslant \log\left(\sum_{i=1}^{N} \exp\left(\frac{x_1}{\varepsilon}\right)\right) \leqslant \frac{\max x_i}{\varepsilon} + \log N$$

ceci est la meme chose que

$$\max x_i \leqslant \varepsilon \log\left(\sum_{i=1}^{N} \exp\left(\frac{x_i}{\varepsilon}\right)\right) \leqslant \max x_i + \varepsilon \log N$$

quand on fait ϵ tend vers l'infini, on a évidement que $\varepsilon \log N$ tend vers 0
Alors on peut conclure que la fonction $f_\epsilon(x) converge\ vers\ max_i x_i\ quand \epsilon\ tend\ vers\ 0$par l'encadrement de l'inégalité

3.4 2.2

Pour ce moment, on peut donc appliquer la résultat précédent pour exprimer cette question d'optimalité sous contraintes comme une question d'optimalité sans contraintes:

$$\max\left\{\sum_i u_i \mu_i + \sum_j v_j \nu_j \mid u, v \in \mathbb{R}^N u_i + v_j \leq C_{ij} \forall (i,j) \in I \times J\right\}$$

ici on remplace $v_j\ par\ \min_i C_{ij} - u_i \quad \forall j \in J$ et puis on remplace $\min_i C_{ij} - u_i \quad \forall j \in J\ par -\varepsilon \log\left(\sum_{i=1}^{N} \exp\left(\frac{-C_{ij}+U_i}{\varepsilon}\right)\right)$ D'apres cela , on peut ecrire notre probleme duale comme un probleme dans une seule variable u sans contrainte :

$$\max\left\{\sum_i u_i \mu_i - \sum_j \left(\varepsilon \log\left(\sum_{i=1}^{N} \exp\left(\frac{-C_{ij} + U_i}{\varepsilon}\right)\right) \nu_j\right)\right\}$$

3.5 2.3

Calculon le gradient de critere:

$$\nabla_{u_i} f_\varepsilon = \mu_i - \sum_j \nu_j \frac{exp(\frac{-C_{ij}+u_i}{\varepsilon})}{\sum_i exp(\frac{-C_{ij}+u_i}{\varepsilon})}$$

4 EXO3

4.1 3.1

On s'intéresse maintenant à la régularisation du problème primal en pénalisant la contrainte de positivité en ajoutant un terme d'entropie on va minimiser le critère avec la terme d'entropie, tout abord, montrons qu'elle est convexe: soit t entre 0 et 1,soit γ , $\beta \in \mathbb{R}_+^{N\times N}$:

$$F(t\gamma + (1-t)\beta) = \sum_{ij} C_{ij}(t\gamma_{ij} + (1-t)\beta_{ij}) + \varepsilon \sum_{j}(t\,(\gamma_{ij} + (1-t)\beta_{ij})\,(\log(t\gamma_{ij} + (1-t)\beta_{ij}) - 1)$$

$$= \sum_{ij} C_{ij}(t\gamma_{ij}+(1-t)\beta_{ij})+\varepsilon(\sum_{ij} t\gamma_{ij}\log(t\gamma_{ij}+(1-t)\beta_{ij})+\sum_{ij}(1-t)\beta_{ij}\log(t\gamma_{ij}+(1-t))\beta_{ij})-\sum_{ij} t\gamma_{ij}-\sum_{ij}(1-t)\beta_{ij}$$

calculons alors $tF(\gamma) + (1-t)F(\beta)$:

$$tF(\gamma)+(1-t)F(\beta) = \sum_{ij} tC_{ij}\gamma_{ij} + \varepsilon\sum_{ij} t\gamma_{ij}(\log(\gamma_{ij}) - 1) + \sum_{ij}(1-t)C_{ij}\beta_{ij} + \varepsilon\sum_{ij}(1-t)\beta_{ij}(\log(\beta_{ij}) - 1)$$

$$= \sum_{ij} C_{ij}(t\gamma_{ij} + (1-t)\beta_{ij}) + \varepsilon(\sum_{ij} t\gamma_{ij}(\log(\gamma_{ij}) + \sum_{ij}(1-t)\beta_{ij}\log(\beta_{ij})) - \sum_{ij} t\gamma_{ij} - \sum_{ij}(1-t)\beta_{ij}$$

en éliminant les parties communs de $F(t\gamma + (1-t)\beta)\ \ et\ \ tF(\gamma) + (1-t)F(\beta)$ il nous suffit de juste comparer

$$\sum_{ij} t\gamma_{ij}\log(\gamma_{ij}) + (1-t)\beta_{ij}\log(\beta_{ij})$$

et

$$\sum_{ij}(t\gamma_{ij} + (1-t)\beta_{ij})\log(t\gamma_{ij} + (1-t)\beta_{ij})$$

dans cette situation ,posons une fonction pour tout x positive, f(x)=xlogx, il nous suffit d'etudier la convexité de cette fonction pour conclure la convexité de notre problème. En fait, f"(x)=1/x ce qui est tout le temps positive , ce qui conclut la convexité de la fonction f, d'après cela on peut trouver facilement que $\sum_{ij} t\gamma_{ij}\log(\gamma_{ij}) + (1-t)\beta_{ij}\log(\beta_{ij})$est plus grand que $\sum_{ij}(t\gamma_{ij} + (1-t)\beta_{ij})\log(t\gamma_{ij} + (1-t)\beta_{ij})$ car chaque terme l'est ,finalement , on a convexité de notre F

et de plus

$$F(\gamma) = \sum_{ij} C_{ij}\gamma_{ij} + \varepsilon\sum_{ij}\gamma_{ij}(\log(\gamma_{ij}) - 1)$$
$$= \sum_{ij}\gamma_{ij}(C_{ij} + \varepsilon\log(\gamma_{ij}) - \varepsilon)$$
$$= \sum_{ij}\varepsilon\gamma_{ij}(\frac{C_{ij}}{\varepsilon} + \log(\gamma_{ij}) - 1)$$
$$= \sum_{ij}\varepsilon\gamma_{ij}(\log(\gamma_{ij}) - \log(exp(-\frac{C_{ij}}{\varepsilon})) - 1)$$
$$= \sum_{ij}\varepsilon\gamma_{ij}(\log(\frac{\gamma_{ij}}{\bar{\gamma_{ij}}}) - 1)$$

avec $\bar{\gamma_{ij}} = exp(\frac{-C_{ij}}{\varepsilon})$

4.2 3.2.1

Cherchons pour ce moment le lagrangien de ce probleme: soit u et v dans R^N les mulitiplicateur de lagrange pour les contraintes

$$L(\gamma*, u, v) = F(\gamma*) + \sum_i u_i(\sum_j \mu_i - \gamma_{ij}*) + \sum_j v_j(\sum_i \nu_j - \gamma_{ij}*)$$

On a alors la condition d'optimalité :

$$\forall i \in I, j \in J. \quad (\nabla L\left(\gamma^*, u, v\right))_{ij}$$

$$= \varepsilon \log\left(\frac{\gamma_{ij}^*}{\bar{\gamma}_{ij}}\right) - u_i - v_j = 0$$

4.3 3.2.2

D'après la question précédente, on peut conclure facilement cette question:

$$\log(\frac{\gamma_{ij}*}{\bar{\gamma_{ij}}}) = \frac{u_i + v_j}{\varepsilon}$$

Ceci implique que:

$$\gamma_{ij}* = a_i b_j \bar{\gamma_{ij}}$$

avec$a_i = exp(u_i/\varepsilon) \;\; et \;\; b_j = exp(v_j/\varepsilon)$

4.4 3.2.3

A partir d'ici, on voudrait savoir les expression de ai et bj, on fait la somme de $\gamma_{ij}*$ respectivement sur i et sur j
quand on fait la somme sur i:

$$\sum_i \gamma_{ij}* = \sum_i a_i b_j \bar{\gamma_{ij}}$$

ce qui implique que

$$\nu_j = \sum_i a_i b_j \bar{\gamma_{ij}}$$

on a alors

$$b_j = \frac{\nu_j}{\sum_i a_i \bar{\gamma_{ij}}}$$

la meme raisonnement pour a_i quand on passe a sommer sur j, on obtient donc:

$$a_i = \frac{\mu_i}{\sum_j b_j \bar{\gamma_{ij}}}$$

4.5 3.2.4

Avec la resultat de la 3.2.3,on est alors capable d'écrire le problème dual correspondant:

$$max \; q(u, v) = F(\gamma*) + \sum_i u_i(\sum_j \mu_i - \gamma_{ij}*) + \sum_j v_j(\sum_i \nu_j - \gamma_{ij}*)$$

en mettant l'expression de $\gamma_{ij}*$

$$= \sum_{ij} \varepsilon a_i b_j \bar{\gamma_{ij}}(\log(a_i b_j) - 1) + \sum_i u_i(\sum_j \mu_i - a_i b_j \bar{\gamma_{ij}}) + \sum_j v_j(\sum_i \nu_j - a_i b_j \bar{\gamma_{ij}})$$

et puis on remplace respectivement les expressions de a_i et b_j,on est capable de faire la simplification, et on obtient donc ce probleme duale:

$$max\ q(u,v) = \sum_i u_i\mu_i + \sum_j v_j\nu_j - \varepsilon \sum_{ij} exp(\frac{u_i + v_j - C_{ij}}{\varepsilon})$$

5 Option A

Dans cette option, on étudie le meme problème d'optimisation, mais on s'étend en dimension 3

5.1 QA.1

Ecrivons tout d'abord le probleme de monge kantorich:

$$\left\{ \begin{array}{l} \min_{\gamma\in\Pi(\rho)} \sum_{i_1,i_2,i_3} C_{i_1,i_2,i_3}\gamma_{i_1,i_2,i_3} \\ \Pi(\mu,\nu,\delta) = \left\{ \begin{array}{l} \sum_{i_1,i_2} \gamma_{i_1,i_2,i_3} = \rho_{i_3} \forall i_3 \in I_3 \\ \sum_{i_1,i_3} \gamma_{i_1,i_2,i_3} = \rho_{i_2} \forall i_2 \in I_2 \\ \sum_{i_2,i_3} \gamma_{i_1,i_2,i_3} = \rho_{i_1} \forall i_1 \in I_1 \end{array} \right\} \end{array} \right\}$$

A l'aide de question suivante, ce qui avec l'expression de $\gamma*$ on peut alors écrire ce dual avec le même raisonnement que la question précédentes, on obtient alors

$$max\ q(u,v,w) = \sum_{i1} u_{i1}\rho_{i1} + \sum_{i2} v_{i2}\rho_{i2} + \sum_{i3} w_{i3}\rho_{i3} - \varepsilon \sum_{i1i2i3} exp(\frac{u_{i1} + v_{i2} + w_{i3} - C_{i1i2i3}}{\varepsilon})$$

5.2 QA.2.1

Comme le meme raisonnement que la question precedente, on peut alors ecrire le probleme de monge kantorich avec la regularisation entropique:

$$\left\{ \begin{array}{l} \min_{\gamma\in\Pi(\rho)} F(\gamma) \\ \Pi(\mu,\nu,\delta) = \left\{ \begin{array}{l} \sum_{i_1,i_2} \gamma_{i_1,i_2,i_3} = \rho_{i_3} \forall i_3 \in I_3 \\ \sum_{i_1,i_3} \gamma_{i_1,i_2,i_3} = \rho_{i_2} \forall i_2 \in I_2 \\ \sum_{i_2,i_3} \gamma_{i_1,i_2,i_3} = \rho_{i_1} \forall i_1 \in I_1 \end{array} \right\} \end{array} \right\}$$

ou $F(\gamma) = \sum_{i_1 i_2 i_3} \varepsilon\gamma_{i_1,i_2,i_3}(\log(\frac{\gamma_{i_1,i_2,i_3}}{\bar{\gamma_{i_1,i_2,i_3}}}) - 1)$, ici $\bar{\gamma_{i_1,i_2,i_3}} = exp(\frac{-C_{i_1 i_2 i_3}}{\varepsilon})$

D'après cette expression, on peut facilement écrire le lagrangien et ainsi les conditions d'optimalité

$$L(\gamma,u,v,w) = F(\gamma*) + \sum_{i1} u_{i1}(\sum_{i2,i3} \rho_{i1} - \gamma_{i_1,i_2,i_3}*) + \sum_{i2} v_{i2}(\sum_{i1,i3} \rho_{i2} - \gamma_{i_1,i_2,i_3}*) + \sum_{i3} u_{i3}(\sum_{i1,i2} \rho_{i3} - \gamma_{i_1,i_2,i_3}*)$$

$$(\nabla_\gamma L(\gamma,u,v,w))_{i_1 i_2 i_3} = \varepsilon \log(\frac{\gamma_{i_1 i_2 i_3}}{\bar{\gamma_{i_1 i_2 i_3}}}) - u_{i1} - v_{i2} - w_{i3} = 0 \quad \forall i \in I$$

5.3 QA.2.2

D'apres la question precedente,on obtient alors que:

$$\gamma_{i_1,i_2,i_3}* = exp(\frac{u_{i1} + v_{i2} + w_{i3}}{\varepsilon})\bar{\gamma_{i_1,i_2,i_3}} = a_{i1}b_{i2}c_{i3}\bar{\gamma_{i_1,i_2,i_3}}$$

ou $a_{i1} = exp(\frac{u_{i1}}{\varepsilon}) \quad b_{i2} = exp(\frac{v_{i2}}{\varepsilon}) \quad c_{i3} = exp(\frac{w_{i3}}{\varepsilon})$

5.4 QA.2.3

Comme le raisonnement de la section 3, quand on fait les sommes de $\gamma_{i1,i2,i3}$ respectivement sur i1,i2,i3 ,on peut obtenir alors:

$$a_{i1} = \frac{\rho_{i1}}{\sum_{i2,i3} b_{i2} c_{i3} \gamma_{i1\bar{i}2i3}}$$

$$b_{i2} = \frac{\rho_{i2}}{\sum_{i1,i3} a_{i1} c_{i3} \gamma_{i1\bar{i}2i3}}$$

$$c_{i3} = \frac{\rho_{i3}}{\sum_{i1,i2} a_{i1} b_{i2} \gamma_{i1\bar{i}2i3}}$$

5.5 QA.2.4

Ecrivons l'algo de Sinkhorn pour faciliter ce problème à l'aide de section 3 et la question précédente:

$$a_{i1}^{n+1} = \frac{\rho_{i1}}{\sum_{i2i3} b_{i2}^{n} c_{i3}^{n} \gamma_{i1\bar{i}2i3}}$$

$$b_{i2}^{n+1} = \frac{\rho_{i2}}{\sum_{i1i3} a_{i1}^{n+1} c_{i3}^{n} \gamma_{i1\bar{i}2i3}}$$

$$c_{i3}^{n+1} = \frac{\rho_{i3}}{\sum_{i1i2} a_{i1}^{n+1} b_{i2}^{n+1} \gamma_{i1\bar{i}2i3}}$$