Dualité Lagrangienne

2 octobre 2024

Lagrangien

Problème d'optimisation convexe

$$\begin{array}{ll} \min_{\mathbf{x}} & f(\mathbf{x}) \\ \text{tel que} & g_i(\mathbf{x}) \leq 0, i = 1, \cdots, k \\ & h_j(\mathbf{x}) = 0, j = 1, \cdots, m \end{array}$$

où f est une fonction convexe et différentiable et g_i , h_j fonctions différentiables définissant des contraintes convexes.

Définition

On définit le Lagrangien comme :

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{k} u_i g_i(\mathbf{x}) + \sum_{j=1}^{m} v_j h_j(\mathbf{x})$$

où $\mathbf{u} \in \mathbb{R}^k$ et $\mathbf{v} \in \mathbb{R}^m$, avec $u_i \geq 0$. Si un des $u_i \leq 0$, on définit implicitement $\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}) = -\infty$

$\mathbf{u}$ et $\mathbf{v}$ sont appelés les multiplicateurs de Lagrange ou variables duales

Propriétés du Lagrangien

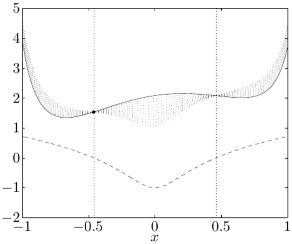
Borne inférieure

pour tout $\mathbf{v}$ et $\mathbf{u} \geq 0$ et $\mathbf{x}$ un point admissible

$$f(\mathbf{x}) \geq \mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v})$$

Preuve

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \sum_{i=1}^{k} \underbrace{u_i g_i(\mathbf{x})}_{\leq 0} + \sum_{j=1}^{m} \underbrace{v_j h_j(\mathbf{x})}_{=0} \leq f(\mathbf{x})$$



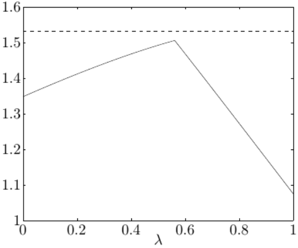
- f : trait plein
- g : trait tiré
- $\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v})$: traits pointillés pour différentes valeurs de $\mathbf{u}$ et $\mathbf{v}$.

Fonction duale de Lagrange

Soit $\mathcal{C}$ l'ensemble des points admissibles du problème original, appelé également problème primal, $f^\star$ la valeur optimale. Minimiser $\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v})$ sur $\mathbf{x}$ donne une borne inférieure sur $f^\star$.

$$f^\star \geq \min_{\mathbf{x} \in \mathcal{C}} \mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}) \geq \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}) := D(\mathbf{u}, \mathbf{v})$$

On appelle $D(\mathbf{u}, \mathbf{v})$ la Fonction duale de Lagrange. Cette fonction donne une borne inf pour tout $\mathbf{u} \geq 0$, $\mathbf{v}$.



- $f^\star$: trait tiret
- λ (notre $\mathbf{u}$)
- $D(\mathbf{u})$: trait plein

Exemple : cas des problèmes quadratiques

le problème primal

$$\begin{array}{ll} \min_{\mathbf{x}\in\mathbb{R}^d} & \frac{1}{2}\mathbf{x}^\top \mathbf{Q}\mathbf{x} + \mathbf{c}^\top \mathbf{x} \\ \text{tel que} & \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0 \end{array}$$

où $\mathbf{Q}$ est symétrique et inversible.

Lagrangien

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}) = \frac{1}{2}\mathbf{x}^\top \mathbf{Q}\mathbf{x} + \mathbf{c}^\top \mathbf{x} + \mathbf{v}^\top(\mathbf{A}\mathbf{x} - \mathbf{b}) - \mathbf{u}^\top \mathbf{x}$$

Fonction duale

$$D(u, v) = \min_{\mathbf{x}\in\mathbb{R}} \mathcal{L}(\mathbf{x}, \mathbf{u}, v) = -\frac{1}{2}(\mathbf{c} - \mathbf{u} + \mathbf{A}^\top \mathbf{v})\mathbf{Q}^{-1}(\mathbf{c} - \mathbf{u} + \mathbf{A}^\top \mathbf{v}) - \mathbf{b}^\top \mathbf{v}$$

Problème dual de Lagrange

Problème primal

$$\begin{array}{ll} \min_{\mathbf{x}} & f(\mathbf{x}) \\ \text{tel que} & g_i(\mathbf{x}) \leq 0, i = 1, \cdots, k \\ & h_j(\mathbf{x}) = 0, j = 1, \cdots, m \end{array}$$

La fonction duale donne une borne inférieure. Donc en maximisant cette borne sur l'ensemble des $\mathbf{u} \geq 0$ et $\mathbf{v}$, on obtient une notre meilleure borne inférieure.

Problème dual

$$\begin{array}{ll} \max_{\mathbf{u},\mathbf{v}} & D(\mathbf{u}, \mathbf{v}) \\ \text{tel que} & \mathbf{u} \geq 0 \end{array}$$

Le problème dual est un problème convexe. Si la valeur optimale $d^\star$ de ce problème existe alors, on a une propriété appelé dualité faible

$$f^\star \geq d^\star$$

Dualité forte

Propriété

La dualité forte signifie que les valeurs optimales des problèmes primal et dual sont telles que

$$f^\star = d^\star$$

maximiser le dual donne alors la meme valeur optimale que minimiser le primal.

Condition de Slater

Si le problème primal est convexe et qu'il existe au moins un point primal $\mathbf{x}$ strictement admissible, alors le probleme admet une dualité forte.

$$\exists \mathbf{x} : \forall i,\ g_i(\mathbf{x}) < 0 \text{ et } \forall j\ h_j(\mathbf{x}) = 0$$

Saut de dualité

Etant donné un $\mathbf{x}$ primal admissible et $\mathbf{u}$ et $\mathbf{v}$ dual admissible, on appelle saut de dualité (duality gap) la quantité

$$f(\mathbf{x}) - D(\mathbf{u}, \mathbf{v})$$

Propriété

On a

$$f(\mathbf{x}) - f^\star \leq f(\mathbf{x}) - D(\mathbf{u}, \mathbf{v})$$

car $f^\star \geq D(\mathbf{u}, \mathbf{v})$. Si le saut de dualité est nul pour un triplet $\mathbf{x}^\star, \mathbf{u}^\star, \mathbf{v}^\star$ alors $\mathbf{x}^\star$ est optimal pour le problème primal et $\mathbf{u}^\star, \mathbf{v}^\star$ l'est pour le problème dual.

Point de vue algorithmique

Le saut de dualité est un critère permettant de suivre l'optimalité : si $f(\mathbf{x}) - D(\mathbf{u}, \mathbf{v}) \leq \epsilon$ alors $f(\mathbf{x}) - f^\star \leq \epsilon$

Conditions de Karush-Kuhn-Tucker

Etant donné le probleme primal

$$\begin{array}{ll} \min_{\mathbf{x}} & f(\mathbf{x}) \\ \text{tel que} & g_i(\mathbf{x}) \leq 0, i = 1, \cdots, k \\ & h_j(\mathbf{x}) = 0, j = 1, \cdots, m \end{array}$$

on définit les conditions KKT comme

Conditions de Karush-Kuhn-Tucker (KKT)

- $\nabla_{\mathbf{x}} f(\mathbf{x}) + \sum_i u_i \nabla_{\mathbf{x}} g_i(\mathbf{x}) + \sum_j v_j \nabla_{\mathbf{x}} h_j(\mathbf{x}) = 0$ — stationarité
- $g_i(\mathbf{x}) \leq 0, h_j(\mathbf{x}) = 0 \forall i = 1, \cdots, k \; j = 1, \cdots, m$ — admissibilité primale
- $u_i \geq 0 \; \forall i$ — admissibilité duale
- $u_i g_i(\mathbf{x}) = 0 \; \forall i$ — complementarité

Conditions d'optimalité

Pour un problème primal avec dualité forte (on peut supposer que les conditions de Slater sont verifiés), on a

$$\mathbf{x}^\star \text{ et } \mathbf{u}^\star \text{ et } \mathbf{v}^\star \text{ solutions des problèmes primales et duales}$$
$$\Leftrightarrow$$
$$\mathbf{x}^\star \text{ et } \mathbf{u}^\star \text{ et } \mathbf{v}^\star \text{ satisfont les conditions KKT.}$$

preuve ($\Leftarrow$)

Supposons $\mathbf{x}^\star$ et $\mathbf{u}^\star$ et $\mathbf{v}^\star$ satisfont les conditions KKT,alors

$$\begin{aligned} D(\mathbf{u}^\star, \mathbf{v}^\star) = & \quad f(\mathbf{x}^\star) + \sum_{i=1}^{k} u_i^\star g_i(\mathbf{x}^\star) + \sum_{j=1}^{m} v_j^\star h_j(\mathbf{x}^\star) \\ = & \quad f(\mathbf{x}^\star) \end{aligned}$$

où la première égalité vient de la définition de $\mathbf{D}(\mathbf{u}, \mathbf{v})$ et la condition de stationarité et la deuxième vient des conditions de complémentarité. Le saut de dualité est nul donc $\mathbf{x}^\star$ et $\mathbf{u}^\star$ et $\mathbf{v}^\star$ sont optimales.

Méthodologie pratique

Utilisation

1. écrire le Lagrangien $\mathcal{L}$ du problème
2. dériver les conditions KKT et la condition de stationnarité
3. si on a une forme analytique de $\mathbf{x}^\star$, dériver la fonction duale et le problème dual
4. résoudre le problème dual
5. en déduire la solution du primal
6. si besoin utiliser les conditions de complémentarités pour déduire la solution