topics in opt notes

libo

2019.8.31

1 kantorovich dual

定义 1.1. 有代价函数 $c(x,y)$, 则 $\varphi(x) = \underset{y}{inf}\, c(x,y) - \varphi(y)$. 共轭函数 $\varphi(x)^* = \underset{y}{sup}\,(x,y) - \varphi(y)$

定理 1.1. 对于 *kantorovich* 问题，它等价与寻找一个费用 $\varphi(x) + \psi(y)$ 使得总费用 $\int \varphi(x) + \psi(y)\, d\pi(x,y)$ 最大，并且这个费用满足 $J(\varphi,\psi) = \varphi(x) + \psi(y) \le c(x,y)$

Proof. $\underset{\pi\in\Pi}{inf} \int c(x,y)\, d\pi(x,y)$ 然后用类似 langrange 乘子法可以验证上式等于 $\underset{\pi}{inf}\, \underset{\varphi,\psi}{sup}\{c(x,y) + \int \varphi dx + \int \psi dy - \int (\varphi + \psi)\, d\pi(x,y)\}$, 交换 inf 和 sup,$\underset{\pi\in\Pi}{inf}\, c(x,y)\, d\pi(x,y) = \underset{\varphi,\psi}{sup}\, \underset{\pi}{inf}\{c(x,y) + \int \varphi dx + \int \psi dy - \int (\varphi+\psi)\, d\pi(x,y)\}$, 在 inf 作用下，若 $\varphi(x) + \psi(y) \neq c(x,y)$, 令 $\pi(x,y) = 0$, 故而上式化为 $\underset{\varphi,\psi}{sup} \int \varphi dx + \int \psi dy$[1] □

定理 1.2. E 是一个线性赋范空间，对于其上的两个凸函数 Θ, Ξ, 那么

$$\underset{x\in E}{inf}\,(\Theta + \Xi) = \underset{z^*\in E^*}{max}\,(-\Theta^*(-z^*) - \Xi^*(z^*))$$

Proof. 令上式左端为 m，考虑两个集合 $C == \{(x,\lambda); \lambda > \Theta(x)\}, C' == \{(y,\mu); \mu \le m - \Xi(y)\}$，那么两个凸集 C, C' 不相交，凸集分离定理告诉我们存在 z^*, α, 使得 $(z^*, x) + \alpha\lambda \le (z^*, y) + \alpha\mu$, 两端同除 α, 化简得到: 存在 $z^*, st, (z^*, x-y) + \Theta + \Xi \le m$ 若令 x=y, 那么我们也可以得到 $\underset{z^*\in E^*}{sup}\ \underset{x,y\in E}{inf}\{\Theta + \Xi + (z^*, x-y)\} \ge m$ □

命题 1.1. 对于 *kantorovich* 对偶中的 φ, ψ, 他们满足 $\varphi = \phi^c, \psi = \phi^{cc}$

Proof. 对于 $sup\, J(\varphi,\psi) = \underset{\varphi,\psi}{sup} \int \varphi dx + \int \psi dy, \varphi + \psi \le c(x,y)$, 若令 $\overline{\varphi}(x) = \psi^c = \underset{y}{inf}\, c(x,y) - \psi(y)$, 因为 $\varphi \le c(x,y) - \psi$, 所以 $\varphi(x) \le \psi^c = \overline{\varphi}$, 即 $J(\overline{\varphi},\psi) \ge J(\varphi,\psi)$, 同理若令 $\overline{\psi} = \overline{\varphi}^c = \underset{x}{inf}\, c(x,y) - \overline{\varphi}(x)$, 有 $J(\overline{\varphi},\overline{\psi}) \ge J(\overline{\varphi},\psi)$, 得证 □

命题 1.2. 若 *cost* 函数等于度量 $d(x)$, 则最优传输等于 $\underset{\phi}{sup} \int \phi d(\mu - \upsilon), ||\phi||_{Lipsz} \le 1$

Proof. 对于 $\underset{\varphi,\psi}{sup} \int \varphi dx + \int \psi dy$, 有 $\varphi = \phi^d, \psi = \phi^{dd}$, 那么 ϕ^d 是 $1 - Lipstz$, 因为 $\phi(y)^d = \underset{x}{inf}\, d(x,y) - \phi(x)$, 那么可以把 $\phi(y)^d - \phi(\overline{y})^d$ 写成 $\underset{x}{inf}\, \underset{\bar{(x)}}{sup}\, d(x,y) - d(\overline{x},\overline{y}) + \varphi(\overline{x}) - \varphi(x)$, 当 $\overline{x} = x$ 的情况知道,$\phi(y)^d - \phi(\overline{y}) \le d(y,\overline{y})$, 同理可知 $\phi(\overline{y})^d - \phi(y) \le d(y,\overline{y})$.
同理，ϕ^{dd} 也是 $1 - Lpstz$, 所以有 $-\phi^d \le \phi^{dd} \le -\phi^d$, 左边是因为 $1 - Lpstz$, 右边可以选择 x=y。 □

2 Briener's

接下来的 cost 函数我们考虑的是欧式距离那么 kantorovich 问题变成 $\underset{\pi\in\Pi}{inf} \int \frac{(x-y)^2}{2} d\pi(x,y) = \int \frac{x^2}{2} dx + \int \frac{y^2}{2} - sup \int xy d\pi(x,y)$, 令 $M_2 = \int \frac{x^2}{2} dx + \int \frac{y^2}{2} dy$, 它的对偶是 $sup \int \varphi dx + \int \psi dy, \varphi(x) + \psi(y) \le \frac{(x-y)^2}{2}$, 我们让 $\tilde{\varphi}(x) = \frac{x^2}{2} - \varphi, \tilde{\psi}(y) = \frac{y^2}{2} - \psi$, 那么 $J(\varphi,\psi) = M_2 - J\left(\tilde{\varphi},\tilde{\psi}\right) = \underset{\pi\in\Pi}{inf} \int \frac{(x-y)^2}{2} d\pi(x,y), \tilde{\varphi} + \tilde{\psi} \ge c(x,y)$, 那么 $inf\, J\left(\tilde{\varphi},\tilde{\psi}\right) = sup \int xy d\pi(x,y)$ 应用上面的命题 1.1, 有 $\tilde{\varphi} = \phi^*, \tilde{\psi} = \phi^{**}$, 我们可以把这两个函数简写 φ, φ^*

命题 2.1. 对于上面的 φ, φ^*,$xy = \varphi(x) + \varphi^*(y) \iff y \in \partial\varphi(x)$[2] $\iff x \in \partial^*(y)$

[1] 证明过程告诉我们传输方案中有很多点的测度是 0
[2] 次微分

Proof. 因为 $xy \le \varphi(x) + \varphi^*(y)$, 所以 $xy \ge \varphi(x) + \varphi^*(y) \iff \forall z \in R^n xy \ge \varphi(x) + yz - \varphi(z) \iff \varphi(z) - \varphi(x) \ge (y, z - x) \iff y \in \partial\varphi(x)$ □

定理 2.1 (Brenier's theorem)**.** 上述最优传输方案 $y(x)$ 是一个凸函数的梯度映射

命题 2.2. 凸函数 φ 的梯度映射 $\nabla\varphi$ 的逆映射是 $\nabla(\varphi^*)$

Proof. $\nabla\varphi^*(x) = \nabla \sup_y xy - \varphi(y) = y, \nabla\varphi(y) = x$, 所以若 $x = \nabla\varphi(x)$ 替换 x，$\nabla\varphi^*(\nabla\varphi(x)) = y, \nabla\varphi(y) = \nabla\varphi(x)$, 所以 y=x □

定理 2.2 (Brenier polar factorization)**.** 对于欧式空间的映射 $h:(W,\lambda) \to (X,\mu)$, 其中 μ 测度是由 h 诱导的测度，那么给定另一欧式空间 (Y,v), 那么 h 可以分解为 $h = \nabla\varphi * s$,s 是 $(W,\lambda) \to (X,v)$ 的保测度映射，φ 是某一凸函数

$\longrightarrow$ 实际上我们对第二节的讨论中 φ, φ^*, 若假设 x,y 的定义域一致, 且测度相等，由于 xy 的对称性，使得 $\varphi = \varphi^*$

定理 2.3. 一个非退化的矩阵存在唯一的极分解[3]

Proof.

$$\begin{aligned}
&\varphi + \varphi = xy = (x, \nabla\varphi(x)) \\
&d(2\varphi) = (dx, \nabla\varphi(x)) + \left(x\nabla^2\varphi dx\right) \\
&\varphi = \int \left(x\nabla^2\varphi dx\right)
\end{aligned}$$

□

[3]极分解不是 QR 分解