No. SHIHE LIU

Date 2301110021.

Construction of the optimal coupling:

Thm: $\mu, \nu \in \mathcal{P}(\mathcal{X})$

$$\|\mu-\nu\|_{TV} = \inf\{\mathbb{P}(X\neq Y) : (X,Y) \text{ is a coupling of } \mu,\nu\}.$$

Proof:

(Integral representation).

Equip $\mathcal{X}$ with measure space structure $(\mathcal{X}, \underset{\text{power set}}{\mathcal{P}(\mathcal{X})}, \mathcal{L})$ with

$$\mathcal{L} = \sum_{x\in\mathcal{X}} \underset{\text{Dirac measure}}{\delta_x} = \# \quad (计数测度)$$

Then view $\mu,\nu \in L^1(\mathcal{X},\mathcal{L})$, denote $dx = d\mathcal{L}(x)$, we have

$$\sum_{x\in A} \mu(x) = \int_A \mu(x)\,dx$$

(Estimate of the upper bound)

Now assume (X,Y) is a coupling of μ,ν with prob distribution (on $\mathcal{X}^2$) $f(x,y)$. Then

$$\int_{\mathcal{X}} f(x,y)\,dy = \mu(x), \quad \int_{\mathcal{X}} f(y,x)\,dy = \nu(x)$$

Hence

$$\|\mu-\nu\|_{TV} = \frac{1}{2}\|\mu-\nu\|_{L^1(\mathcal{X},\mathcal{L})}$$

$$= \frac{1}{2}\int_{\mathcal{X}} \left|\int_{\mathcal{X}} f(x,y)\,dy - \int_{\mathcal{X}} f(y,x)\,dy\right| dx$$

① $$\le \frac{1}{2}\int_{\mathcal{X}}\int_{\mathcal{X}} \underbrace{|f(x,y)-f(y,x)|}_{\text{vanish at diagonal } x=y}\,dy\,dx$$

$$= \frac{1}{2}\int_{\mathcal{X}^2\setminus\{x=y\}} |f(x,y)-f(y,x)|\,dx\,dy$$

No.

Date

$$② \le \frac{1}{2}\int_{\mathcal{X}^2} 1_{\{x\neq y\}} |f(x,y)| + |f(y,x)| \, dx dy$$

by symmetricity of x, y, Fubini

$$= \int_{\mathcal{X}^2} 1_{\{x\neq y\}} |f(x,y)| \, dx dy$$

$$= \mathbb{P}(X \neq Y).$$, take the inf of (X, Y).

(Geometric construction of optimal coupling).

The optimal coupling will ensure the inequiv ①, ② take equiv.

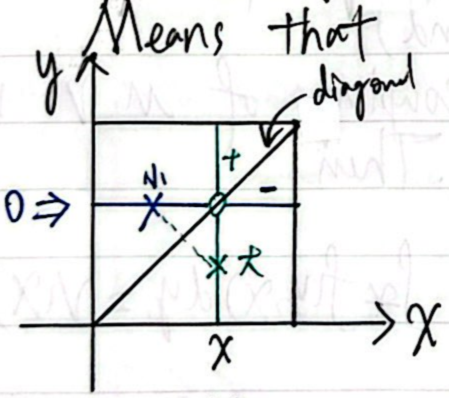
① $\Rightarrow$ for $\forall$ fixed x, either

$$f(x,y) \ge f(y,x), \quad \forall y \in \mathcal{X}$$

or

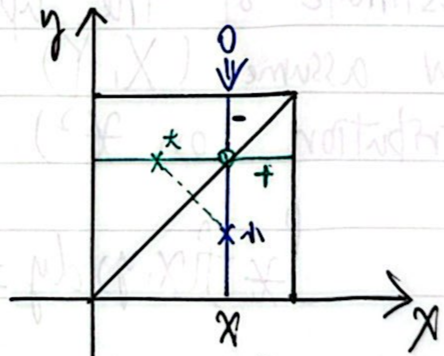
$$f(x,y) \le f(y,x), \quad \forall y \in \mathcal{X}$$

Means that



(A)

or



(B)

② $\Rightarrow$ $|f(x,y) - f(y,x)| = |f(x,y)| + |f(y,x)|, \ \forall x \neq y$

Due to $f \ge 0$

$$\Rightarrow \min\{f(x,y), f(y,x)\} = 0.$$

Hence the minimal part (blue line) in above pictures will be zero at off-diagonal place.

So how to judge whether (A) or (B)? By the

No.

Date

definition of coupling, notice that for fixed choosen $x \in \mathcal{X}$

$$\text{sum of column} = \mu(x)$$
$$\text{sum of row} = \nu(x).$$

Hence only to judge

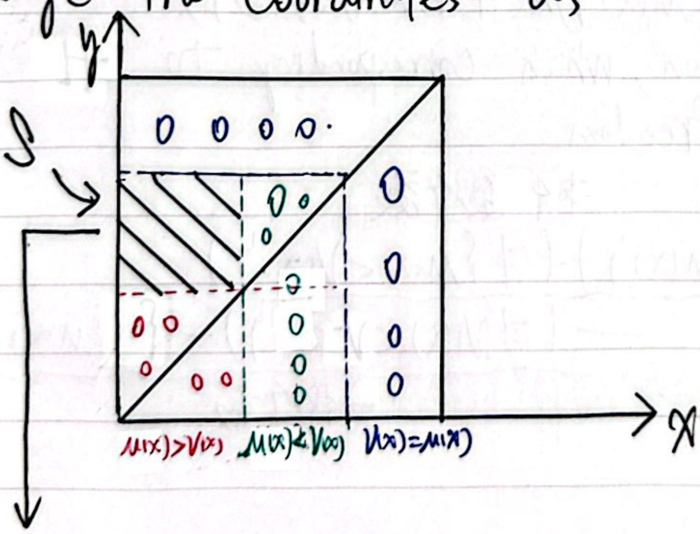
"$\mu(x) > \nu(x)$" $\overset{①}{\Longrightarrow}$ (A) $\overset{②}{\Longrightarrow}$ diagonal $f(x,x) = \nu(x)$.
off-diag row all zero.

"$\mu(x) < \nu(x)$" $\overset{①}{\Longrightarrow}$ (B) $\overset{②}{\Longrightarrow}$ diagonal $f(x,x) = \mu(x)$
off-diag column all zero

"$\mu(x) = \nu(x)$" $\Longrightarrow$ diagonal $f(x,x) = \nu(x) = \mu(x)$
both off-diag column & row all zer

$$\Rightarrow \quad f(x,x) = \mu(x) \wedge \nu(x).$$

Rearrange the coordintes as



In this sequare we can easily get

$$\sum_y f(x,y) = \mu(x) - \nu(x) > 0$$

$$\sum_x f(x,y) = \nu(y) - \mu(y) > 0.$$

No.

Date

and totally $\sum_{(x,y)\in S} f(x,y) = \sum_x (\mu(x)-\nu(x))$

$W \sim \mathrm{Bern}(p)$

$$= \|\mu-\nu\|_{TV}$$

$$\Rightarrow \mathbb{P}(X=Y) = \sum_{x\in\mathcal{X}} \mu(x)\wedge\nu(x) = 1-\|\mu-\nu\|_{TV} = p$$

Nomolize $f|_S / \|\mu-\nu\|_{TV}$ to be a prob. distri on S, with is the coupling of

$$(\mu(x)-\nu(x))/\|\mu-\nu\|_{TV} = \phi_1(x)$$

$$(\nu(y)-\mu(y))/\|\mu-\nu\|_{TV} = \phi_2(y)$$

And if we take the simpliest independent coupling

$$f|_S(x,y)/\|\mu-\nu\|_{TV} = \phi_1(x)\,\phi_2(y).$$

Then $f(x,y)$ is the distribution of what we constructed in the book. What's more, to compute the dimension of optimal coupling, it's equal to a

$$(\#\{\nu(x)<\mu(x)\})\times(\#\{\mu(x)<\nu(x)\})$$

size coupling's dimension, which corresponding to a liner equation system with freedom

$$F \geq \underbrace{(\#\{\nu(x)<\mu(x)\})\cdot(\#\{\mu(x)<\nu(x)\})}_{\text{未知数个数}} - \underbrace{(\#\{\nu(x)<\mu(x)\})-(\#\{\mu(x)<\nu(x)\})}_{\text{方程个数}}.$$

#

What will happens if $\mathcal{X}$ is not discrete? For example $\mathcal{X}=\mathbb{R}$ with $\mathcal{L}=\mathcal{L}^1$ on $\mathbb{R}$?

Conjugation:

$$\|\mu-\nu\|_{TV} \leq \lim_{\delta\to 0+} \inf\{\mathbb{P}(|X-Y|>\delta) \mid (X,Y) \text{ is coupling of } \mu,\nu\}.$$

Campus