TVD contraction

Let P be a Markov kernel with stationary distr^n π.

Then $\|P^{n+1}(x,\cdot) - \pi(\cdot)\|_{TV} \le \|P^n(x,\cdot) - \pi(\cdot)\|_{TV}$

For any $A \in \Omega$, $|P^{n+1}(x,A) - \pi(A)| = \left|\int_{y\in\mathcal{X}} P^n(x,dy)\,P(y,A) - \int_{y\in\mathcal{X}} \pi(dy)\,P(y,A)\right|$

$= \left|\int_{y\in\mathcal{X}} P^n(x,dy)\,h(y) - \int_{y\in\mathcal{X}} \pi(dy)\,h(y)\right| \le \|P^n(x,\cdot) - \pi(\cdot)\|$ by (i)

where $h(y) = P(y,A) : \mathcal{X} \to [0,1]$. Take $\sup_{A\in\Omega}$.

W_1 equivalence

Let $W_1^C(\mu,\nu) = W_1(\mu,\nu) = \inf_{\gamma\in\Gamma} \mathbb{E}_\gamma[D(X,Y)]$.

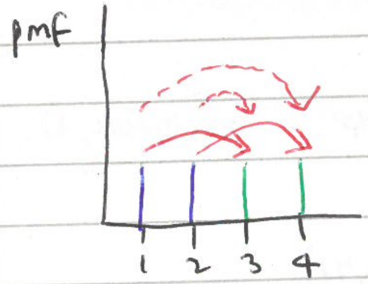
$W_1^L(\mu,\nu) = \sup_{h\in\mathcal{L}} \int h(x)\,d\mu(x) - \int h(x)\,d\nu = \sup_{h\in\mathcal{L}} \mathbb{E}_\mu[h(X)] - \mathbb{E}_\nu[h(Y)]$.

Then $\underline{W_1^C \equiv W_1^L}$. Here $\mathcal{L}$ = set of Lipschitz-1 functions $= \{h : \mathcal{X} \to \mathbb{R} : |h(x) - h(y)| \le D(x,y)\}$.

(4) $W_1^L \le W_1^C$. Let $\gamma_* = \arg\inf_{\gamma\in\Gamma} \mathbb{E}_\gamma[D(X,Y)]$. Let $h_* = \arg\sup_{h\in\mathcal{L}} \mathbb{E}_\mu[h] - \mathbb{E}_\nu[h]$.

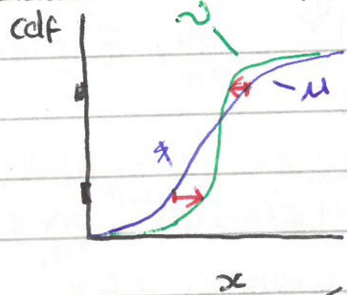
Then $W_1^L(\mu,\nu) = \int [h(x) - h(y)]\,\gamma(dx,dy) \le \int D(x,y)\,\gamma(dx,dy) = W_1^C(\mu,\nu)$.

But why is the inequality tight?



To minimise $\mathbb{E}_\gamma[D^2]$ we need the → map, $(W_2 = 2)$

Could also minimise $\mathbb{E}_\gamma[D]$ using ⇢ $(W_1 = 2)$



Consider the infinitesimal chunk of mass ▪. To shift from μ to ν this should shift according to ↦

switches direction when they cross.

So $\mathbb{E}_\mu[h(x)] - \mathbb{E}_\nu[h(x)]$ for h above

Now $\underset{(\text{any } h)}{\mathbb{E}_\mu[h(X)]} = \int_{\mathcal{X}} h(x)\,d\mu(x) = \int_0^1 h(F^{-1}(p))\,dp$ $\approx \int$ mass × how far it moves.

Transport plan

	.45	4			.15	.3
	.05	3			.05	
μ	.2	2	.1	.1		
	.3	1	.3			
			1	2	3	4
			.4	.1	.2	.3 ν

Cost $= .15 \times (4-3) + .1 \times (2-1)$

ie sum of mass × how far it moves.