OPTimal Transport plan

Fecha:

$$\min \int_{X \times Y} c(x,y)\, d(x,y) \quad , \quad d(x,y) = d(y|x)\,dx$$

s.a. $X \sim \mu_{ref}$ $\qquad dy = \int_X d(y|x)\,dx$

 $Y \sim \mu_{tar}$

$$Y = T(x) \quad , \quad d(y|x) = \frac{dy}{dx} = |\nabla T(x)|$$

$$\min \int_X c(x, T(x))\, d\mu_{ref}(x)$$

s.a. $\mu_{tar}(\cdot) = \mu_{ref}(T^{-1}(\cdot))$

$\mu_{tar} = T_\# \mu_{ref}$

$\pi = T_\# \eta = \eta \circ T^{-1} |\det \nabla T^{-1}|$

$$\min \int_X c(x, T(x))\, d\eta + \lambda D_{KL}\left(\eta \circ T^{-1} |\det \nabla T^{-1}| \,\Big\|\, \pi\right)$$

$$D_{KL}(T^{-1}_\# \pi \| \eta) = D_{KL}(\pi \| T_\# \eta) = \int \pi(x) \ln \frac{\pi(x)}{T_\# \eta(x)}\, dx = \int T^{-1}_\# \pi(x) \ln \frac{T^{-1}_\# \pi}{\eta}$$

$$D_{KL}(\eta \| T^{-1}_\# \pi) = D_{KL}(T_\# \eta \| \pi) = \int \eta(x) \ln \frac{\eta(x)}{T^{-1}_\# \pi(x)}\, dx = \int T_\# \eta(x) \ln \frac{T_\# \eta}{\pi(x)}$$

Fecha:

La divergencia D_{KL} se podría escoger como otra función, al igual que la función de costo $c(x,T(x))$

En machine learning, no tenemos acceso a π directamente, sino a muestras $\{y_i\}_{i=1}^{m} \sim \pi$

$$D_{KL}(\pi \| T_\# \eta) = \int \pi(x) \ln \frac{\pi(x)}{T_\# \eta(x)} dx = \int \pi(x) \ln \pi(x) dx - \int \pi(x) \ln T_\# \eta(x) dx$$

$$= \underbrace{H(\pi)}_{C} - \underbrace{\int \pi(x) \ln T_\# \eta(x) dx}_{\approx \frac{1}{m} \sum_{i=1}^{m} \ln T_\# \eta(x_i)}$$

$$\min \underbrace{\int c(x, T(x)) d\eta}_{} - \frac{\lambda}{m} \sum_{i=1}^{m} \ln\left(\eta \circ T^{-1}(y_i) |\nabla T^{-1}(y_i)|\right)$$

$$\int c(x,T(x)) \eta(x) dx = \int \|x - T(x)\|^2 \eta(x) dx$$

$$= \int \|T^{-1}(y) - y\|^2 \underbrace{\eta \circ T^{-1}(y) |\nabla T^{-1}(y)|}_{\pi(y)} dy$$

$$\approx \frac{1}{m} \sum_{i=1}^{m} \|T^{-1}(y_i) - y_i\|^2$$

Fecha:

$$\min \frac{1}{n} \sum_{i=1}^{m} \| T^{-1}(y_i) - y_i \|^2 - \lambda \ln \left[\eta \circ T^{-1}(y_i) \, |\nabla T^{-1}(y_i)| \right] \quad \lambda > 0$$

$$\Leftrightarrow$$

$$\min \frac{1}{n} \sum_{i=1}^{m} \varepsilon \| T^{-1}(y_i) - y_i \|^2 - \ln \left(\eta \circ T^{-1}(y_i) \, |\nabla T^{-1}(y_i)| \right) \quad \varepsilon > 0$$

$$x \to +\infty \, , \quad \varepsilon \to 0^+ \quad \lambda = \frac{1}{\varepsilon}$$

$$T_\alpha : X \to Y$$

$$T : \alpha \times X \to Y \qquad T^{-1} : \alpha \times Y \to X$$

$$\alpha \sim \Gamma(\cdot)$$

$$P(y) = \int P_\alpha(x) \, d\alpha = \int \eta \circ T_\alpha^{-1}(y) \, |\nabla T_\alpha^{-1}(y)| \, d\alpha$$

$$\int c(x,y) \, d(x,y) = \iint c(x, T_\alpha(x)) \, d\alpha \, dx$$

$$dX = T_\alpha(x) \, d\alpha \, , \quad dy|_\alpha = T_\alpha(x) dx$$

$$dy|x = T_\alpha(x) d\alpha$$

$$d(x,y) = d(x) \, d(y|x) \qquad N(\alpha x | 0, K) \, d\Gamma^{-1}(\alpha) = T(\nu, \rho,$$

$$= d(y|x) \, dx$$

$$= d(T_\alpha(x)) \, d\alpha \, dx$$

$$= d(T_\alpha(x)) \, dx \, d\alpha$$

Fecha:

$X \sim \mu_{ref}$, $\mu_{ref}(A) = \int_A \eta(x)\,dx$

$Y \sim \mu_{tar}$, $\mu_{tar}(A) = \int_A \pi(x)\,dx$ $X \xrightarrow{T} Y$

$T(x) \overset{d}{=} Y$, $T_\# \mu_{ref} = \mu_{tar}$, $\mu_{ref}(T^{-1}(A)) = \mu_{tar}(A)$

$$\mu_{tar}(A) = \int_A \pi(x)\,dx = \int_{T^{-1}(A)} \eta(x)\,dx = \int_A \eta(T^{-1}(x)) \cdot |\nabla T^{-1}(x)|\,dx$$

$$\Rightarrow \pi(x) = \eta(T^{-1}(x))\,|\nabla T^{-1}(x)| \quad , \quad S = T^{-1}$$

$$= \eta(S(x))\,|\nabla S(x)|$$

$$\int g(x)\pi(x)\,dx = \int g(x)\,\eta(S(x))\,|\nabla S(x)|\,dx$$

$$= \int g(T(x))\,\eta(x)\,dx$$

$T_\# \eta = \pi$, $S_\# \pi = \eta$

Consideremos una fuente aleatoria independiente de X.

$\alpha \sim p(\cdot)$ tal que $\mu_\alpha(A) = \int_A p(\alpha)\,d\alpha$, sean ~~mapas~~ Transport

$T : \alpha \times X \to Y$ $T(\alpha, S(\alpha, y)) = y$

$S : \alpha \times Y \to X$ $S(\alpha, T(\alpha, x)) = x$

Fecha:

$x \sim \eta(\cdot)$ $\quad \Rightarrow y = T(\alpha, x) \sim \overline{\Pi}(\cdot)$

$\alpha \sim \Gamma(\cdot)$ $\quad x = S(\alpha, y) \sim \eta(\cdot)$

$x \perp \alpha$

$$\mu_{ref}(A) = \int_A \eta(x)\Gamma(\alpha)\,dx\,d\alpha = \int_A \eta(\alpha, x)\,d(\alpha,x)$$

$$\mu_{ref}^{x}(B) = \mu_{ref}(\mathbb{R}\times B) = \int_{\mathbb{R}\times B} \eta(\alpha, x)\,d(\alpha,x)$$

$$= \int_{\mathbb{R}} \underbrace{\int_B \eta(\alpha, x)\,d(x|\alpha)}_{\mu_{ref}^{y}(B|\alpha) = \mu_{ref}^{y|\alpha}(B)}\,d(\alpha) = \mathbb{E}_\alpha\left[\mu_{ref}^{x|\alpha}(B)\right]$$

$$= \int_B \int_{\mathbb{R}} \eta(\alpha, x)\,d(\alpha|x)\,dx = \int_B \eta^x(x)\,dx$$

En el caso que $x \perp \alpha$, $d(\alpha|x) = d\alpha$, $\alpha \sim \eta(\alpha)$

$$\eta^x(x) = \int_{\mathbb{R}} \eta(\alpha, x)\,d\alpha = \mathbb{E}_\alpha\left[\eta(\alpha, x)\right]$$

$$\mu_{tar}^{y}(B) = \int_B \int_{\mathbb{R}} \eta(\alpha, S(\alpha, x))\,|\nabla_x S(\alpha, x)|\,\Gamma(\alpha)\,d\alpha\,dx$$

$$\overline{\Pi}^{y}(y) = \int_{\mathbb{R}} \eta(\alpha, S(\alpha, x))\,|\nabla_x S(\alpha, x)|\,\Gamma(\alpha)\,d\alpha$$

Fecha:

$\cancel{d\gamma(x,y)} =$ $\inf \int_{X \times Y} c(x,y)\, d\gamma(x,y)$

s.a. $\gamma \in \Gamma(\mu, \nu)$

$$\int_X d\gamma(x,y) = \nu = \eta^x_{ref}$$

$$\int_Y d\gamma(x,y) = \mu = \pi^y_{tar}$$

$$d\gamma(\alpha, x, y) = \eta_{ref}(x) \cdot \Gamma(\alpha)\, \delta_{T(\alpha,x)=y}$$

$$d\gamma(x,y) = \int_{\mathbb{R}} d\gamma(\alpha, x, y) = \int_{\mathbb{R}} \eta(x)\, \Gamma(\alpha)\, \delta_{T(\alpha,x)=y}\, d\alpha$$

$$= \eta(x) \int_{\{\alpha \mid T(\alpha,x)=y\}} \Gamma(\alpha)\, d\alpha = \eta(x) \cdot \Gamma(y \mid x)$$

$$\int_{X \times Y} c(x,y)\, \eta(x) \int_{\{\alpha \mid T(\alpha,x)=y\}} \Gamma(\alpha)\, d\alpha \; dx \;\; \rangle$$

$$\int_{X \times Y} \int_{\{\alpha \mid T(\alpha,x)=y\}} c(x, T(\alpha,x))\, \eta(x)\, \Gamma(\alpha)\, d\alpha\, dx$$

Fecha:

$$\int_{X\times Y} c(x,y)\,\eta(x)\,\Gamma(y|x)\,dy\,dx$$

$$= \int_{X\times\mathbb{R}} c(x,T(z,x))\,\eta(x)\,\Gamma(z)\,dz\,dx$$

$$\boxed{\min_F \int_{X\times\mathbb{R}} c(x,T(z,x))\,\eta(x)\,\Gamma(z)\,dz\,dx + \lambda D_{KL}\left[\int_{\mathbb{R}} \eta(z,S(z,x))\,|\nabla_x S(z,x)|\,\Gamma(z)\,dz \,\Big\|\, \pi(x)\right]}$$

en el contexto de GPs:

$\eta(x) = \mathcal{N}(\mu_x, \Sigma_x)$ medida de referencia Gaussiana u otra, con acceso a la logP y puede samplear

$\{X_i\}_{i=1}^{m} \sim \pi(x)$ observaciones de la medida objetivo

$\Gamma(z)$ distribución con acceso a la logP y samplear fácil.

$T(z,x)$ evaluable fácil

$S(z,y)$ derivable fácil

Costo Trasporte

Fecha:

(1) $$\int_{X\times\mathbb{R}} c(x, T(\alpha,x))\,\eta(x)\,\Gamma(\alpha)\,d\alpha\,dx = \int_X \hat{c}(x,\hat{T}(x))\,\eta(x)\,dx$$

(2) $$= \int_{Y\times\mathbb{R}} c(S(\alpha,y), y)\,\Pi(y)\,\Gamma(\alpha)\,dy\,d\alpha = \int_Y \hat{c}(\hat{S}(y),y)\,\Pi(y)\,dy$$

(1) y (2) son equivalentes, pero si no es posible resolver la integral de forma exacta,

(1) $$\approx \frac{1}{m}\sum_{i=1}^{m}\int_{\mathbb{R}} c(x_i, T(\alpha,x_i))\,\Gamma(\alpha)\,d\alpha \quad \text{con } \{x_i\}_{i=1}^{m}\sim\eta(\cdot)$$

$$\approx \frac{1}{m}\sum_{i=1}^{m} c(x_i, T(\alpha_i,x_i)) \quad \text{con } \{\alpha_i,x_i\}_{i=1}^{m}\sim(\Gamma(\cdot),\eta(\cdot))$$

con $\{y_i\}_{i=1}^{m}\sim\Pi(y)$ de forma

(2) $$\approx \frac{1}{n}\sum_{i=1}^{n}\int_{\mathbb{R}} c(S(\alpha,y_i),y_i)\,\Gamma(\alpha)\,d\alpha \quad \text{con } \{y_i\}_{i=1}^{n}\sim\Pi(\cdot)$$

$$\approx \frac{1}{n}\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m} c(S(\alpha_j,y_i),y_i) \quad \text{con } \{\alpha_j\}_{j=1}^{m}\sim\Gamma(\cdot)$$

Fecha:

Restricción

$$T_\# \mu_{ref} = \mu_{tar} \iff D_{KL}(T_\# \mu_{ref} \| \mu_{tar}) = D_{KL}(\mu_{tar} \| T_\# \mu_{ref}) = 0$$

Se puede usar cualquiera de las 2 direcciones u otra medida de divergencia pero se usa D_{KL} por la siguiente propiedad:

① $D_{KL}(S_\# \overline{\pi} \| \eta) = D_{KL}(\overline{\pi} \| T_\# \eta) = \int \overline{\pi}(x) \ln \frac{\overline{\pi}(x)}{T_\# \eta(x)} dx = \int S_\# \overline{\pi}(x) \ln \frac{S_\# \overline{\pi}(x)}{\eta(x)}$

② $D_{KL}(\eta \| S_\# \overline{\pi}) = D_{KL}(T_\# \eta \| \overline{\pi}) = \int \eta(x) \ln \frac{\eta(x)}{S_\# \overline{\pi}(x)} dx = \int T_\# \eta(x) \ln \frac{T_\# \eta(x)}{\overline{\pi}(x)}$

① $= \int \overline{\pi} \ln \overline{\pi} - \int \overline{\pi} \ln T_\# \eta = \int S_\# \overline{\pi} \ln S_\# \overline{\pi} - \int S_\# \overline{\pi} \ln \eta$

② $= \int \eta \ln \eta - \int \eta \ln S_\# \overline{\pi} = \int T_\# \eta \ln T_\# \eta - \int T_\# \eta \ln \overline{\pi}$

en todas es necesario evaluar $\overline{\pi}(x)$ o $S_\# \overline{\pi}(x)$, pero la expresión ① cumple que, para el problema de optimización de η y T, el término $\int \overline{\pi} \ln \overline{\pi}$ es constante, por lo que se puede replantear el problema como

Fecha:

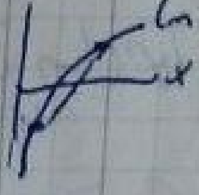
$$\inf_{\eta,T}^{\lambda} \int_{X\times\Gamma} C(x,T(\alpha,x))\,\eta(x)\,\Gamma(\alpha)\,dx\,d\alpha - \lambda\int \Pi(y)\ln T_\#\eta(y)\,dy$$

Con $\{y_i\}_{i=1}^{n} \sim \Pi(\cdot)$

$$\int \Pi(y)\ln T_\#\eta(y)\,dy \approx \frac{1}{n}\sum_{i=1}^{n} \ln T_\#\eta(y_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \ln \int_\Gamma \eta(\alpha, S(\alpha,x_i))\,|\nabla_x S(\alpha,x_i)|\,\Gamma(\alpha)\,d\alpha$$

$$\approx \frac{1}{n}\sum_{i=1}^{n} \ln \frac{1}{m}\sum_{j=1}^{m} \eta(\alpha_j, S(\alpha_j,x_i))\,|\nabla_x S(\alpha_j,x_i)|$$ Con $\{\alpha_j\}_{j=1}^{m} \sim \Gamma(\cdot)$



$\mathbb{E}(\ln(x)) \le \ln \mathbb{E}(x) = \ln 1.5$

$\int_1^2 \ln(x)\,dx = x\ln x - x\Big|_1^2 = 2\ln 2 - 2 + 1 = \ln 4 - \ln e = \ln\frac{4}{e}$

$$\ge \frac{1}{n}\frac{1}{m}\sum_{i=1}^{n}\sum_{j=1}^{m} \ln\left[\eta(\alpha_j, S(\alpha_j,x_i))\,|\nabla_x S(\alpha_j,x_i)|\right]$$

$$\inf_{\eta,T}^{\varepsilon}\ \varepsilon\int_{X\times\Gamma} C(x,T(\alpha,x))\,\eta(x)\,\Gamma(\alpha)\,dx\,d\alpha - \int \Pi(y)\ln T_\#\eta\,dy$$

Fecha:

Caso 1 Caso trivial: $C(x,y)=0$, $T = I_d$, $\Gamma(\alpha)=\delta_0$

$\eta(x) = GP(m(t), k(t,\tilde{t}))$

$Y_{obs} = \{(t_i, y_i)\}_{i=1}^{n} = (\vec{t}, \vec{y})$

$$\inf_{m(t),\, k(t,\tilde{t})} -\ln GP(\vec{y} \mid m(\vec{t}), k(\vec{t},\vec{t}))$$

$$= -\frac{n}{2}\ln 2\pi - \frac{1}{2}(\vec{y}-m(\vec{t}))^T k(\vec{t},\vec{t})^{-1}(\vec{y}-m(\vec{t})) - \frac{1}{2}\ln|k(\vec{t},\vec{t})|$$

NLL en GP standard//

Caso 2 $T = \begin{bmatrix}\varphi(x_1) \\ \vdots \\ \varphi(x_n)\end{bmatrix} \Rightarrow$

$$\inf_{m(t),\, k(t,t),\, \varphi} -\frac{n}{2}\ln 2\pi - \frac{1}{2}(\varphi^{-1}(\vec{y})-m(\vec{t}))\, k(\vec{t},\vec{t})^{-1}(\varphi^{-1}(\vec{y})-m(\vec{t})) - \frac{1}{2}\ln|k(\vec{t},\vec{t})|$$

$$\underbrace{+ \ln|\nabla \vec{T}^{-1}(\vec{y})|}_{= \sum_{i=1}^{n}\ln\left(\frac{d\varphi^{-1}}{dy}(y_i)\right)}$$

NLL en WGP standard

Caso 3 $T(\alpha, x) = \sqrt{\alpha}\cdot X$, $\Gamma(\alpha) = Gamma^{-1}(\nu/2, \frac{\nu-2}{2})$

$\eta(x) = GP(0, k(t,\tilde{t}))$

$$\inf_{\nu,\, k} -\ln TP(\vec{y} \mid \nu, 0, k(\vec{t},\vec{t}))$$

NLL en Student-t Process//

Fecha:

Caso 4 Sea $\eta(\cdot) = GP(0, \delta_{t=\tilde{t}}) = WNP()$

$$T(x) = \Sigma^{1/2} x \quad \text{donde} \quad \Sigma = K(t,\tilde{t})$$
$$T^{-1}(y) = \Sigma^{-1/2} y$$

$$\inf_K -\frac{n}{2}\ln 2\pi - \frac{1}{2}(\Sigma^{-1/2}y)^T(\Sigma^{-1/2}y) - \cancel{\frac{1}{2}\ln|I|}^{\,0}$$
$$+ \ln|\Sigma^{-1/2}|$$

$$= \inf_K -\frac{n}{2}\ln 2\pi - \frac{1}{2} y^T \Sigma^{-1} y - \frac{1}{2}\ln|\Sigma|$$

NLL de GP standard, pero donde $\eta(\cdot)$ es fija sin parámetros y T depende del ~~#~~ índice t.

Caso 5 Si $c(x,y) = \|x-y\|^2$

$$\int_X \|x - \Sigma^{1/2}x\|^2 \eta(x)\,dx > 0 \quad \eta(x) = GP(0, I)$$

vs $$\int_X \|x - \cancel{x}\|^2 \eta_K(x)\,dx = 0, \quad \eta_K(x) = GP(0, K)$$

$$\left.\begin{array}{l} T(x) = \begin{pmatrix} m(t_1) + \sigma(t_1)\cdot x_1 \\ \vdots \\ m(t_n) + \sigma(t_n)\cdot x_n \end{pmatrix} \\ \\ x \sim GP(0, I) \end{array}\right\} T(x) \sim GP(m(t), \sigma(t))$$

Fecha:

$$\int_X \|x - T(x)\|^2 \eta(x)\,dx = \sum_{i=1}^{m} \int_{\mathbb{R}} (x - m(t_i) - \sigma(t_i)\cdot x)^2 N(x|0,1)\,dx$$

$$(m(t) + \sigma(t)\cdot x - x)^2 = (m(t) + (\sigma(t)-1)\cdot x)^2$$

$$= m^2(t) + (\sigma(t)-1)^2\cdot x^2 + 2m(t)(\sigma(t)-1)\cdot x \qquad / \int N(x)dx$$

$$= m^2(t) + (\sigma(t)-1)^2$$

$$\inf_{m,\sigma} \varepsilon \sum_{i=1}^{m} m^2(t_i) + (\sigma(t_i)-1)^2 - \frac{n}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_i - m(t_i)}{\sigma(t_i)}\right)^2$$

$$- \frac{1}{2}\sum_{i=1}^{n} \ln \sigma^2(t_i)$$

def sea $r(t,\bar{t})$ un kernel de correlación, ie. es un kernel de covarianza y cumple que

$-1 \leq r(t,\bar{t}) \leq 1 \quad \forall t,\bar{t}$ — todo kernel $K(t,\bar{t})$

$r(t,t) = 1 \quad \forall t$ — se puede expresar como

$$K(t,\bar{t}) = \sigma(t)\sigma(\bar{t})\,r(t,\bar{t}), \quad \text{donde } \sigma(t) = \sqrt{K(t,t)}$$

PROP $m(t), \sigma(t)$ definen la marginal del proceso

$r(t,\bar{t})$ define la cópula (dependencia) del proceso

Fecha:

$$T_K(X) = \Sigma^{1/2} X \quad \text{con} \quad \Sigma = K(\vec{t}, \vec{t})$$

$$\Sigma = \Sigma^{T/2} \Sigma^{1/2}$$

$$(\Sigma)_{ij} = \sigma(t_i)\sigma(t_j)r(t_i, t_j) \qquad R = r(\vec{t}, \vec{t})$$

$$\Gamma = \begin{bmatrix} \sigma(t_1) & & 0 \\ & \ddots & \\ 0 & & \sigma(t_n) \end{bmatrix}$$

$$\Sigma = \Gamma \times R \times \Gamma$$

$$= (\Gamma R^{T/2})(R^{1/2}\Gamma) = (R^{1/2}\Gamma)^T (R^{1/2}\Gamma)$$

$$T_K(X) = \Sigma^{1/2} X = (R^{1/2}\Gamma)X = R^{1/2}(\Gamma X)$$

$$= T_r \circ T_\sigma X = T_\sigma \circ T_r X = \Gamma(R^{1/2} X)$$

$$T_m(X) = m(\vec{t}) + X$$

$$T_{GP}(X) = T_m \circ T_r \circ T_\sigma(X) = T_m \circ T_\sigma \circ T_r(X)$$

$$T_m \circ T_\sigma(X) = \begin{bmatrix} m(t_1) + \sigma(t_1)x_1 \\ \vdots \\ \dots \end{bmatrix}$$

$$T_{TP}(X) = T_m \circ T_\sigma \circ T_\nu \circ T_r(X)$$

$$T_\nu(X) = \sqrt{r} X \qquad r \sim \Gamma^{-1}\left(\nu/2, \frac{\nu - 2}{2}\right)$$

Fecha:

$$T_{WTP}(x) = \underbrace{T_{\varphi} \circ T_{m} \circ T_{\sigma}}_{\text{marginal}} \circ \underbrace{T_{v} \circ T_{r}}_{\text{copula}}(x)$$

$$T_{\varphi}(x) = \begin{bmatrix} \varphi(x_1) \\ \vdots \\ \varphi(x_n) \end{bmatrix}$$

$$S_{WTP}(y) = S_{r} \circ S_{v} \circ S_{\sigma} \circ S_{m} \circ S_{\varphi}(y)$$

1. Mirar ejemplo Steer
2. Mirar otros trasformar (GPmm)
3. Pensar a otros bayes ad-hoc, el problema estadístico o numérico tiene ~~[illegible]~~ corto natural
4. Familias de transportes que sea el problema ad-hoc, hasta donde se pueda llegar
5. A términos de modelo es convexo, al parametrizar cago la convexidad, al minimizar cada modelo ¿puede generar a partir de sus iteraciones una distribución mejor? Bayesiano, nodelone
6. Flujo gradiente otto kinder 98

Pensar el problema como un algoritmo de gradiente en el espacio de medida (W^2 o algún otro).

An

(7) 

ver capacidad de la transformación, según la cantidad de fuentes aleatorias. transformación por coordenadas, aumentando la dimensión de las fuentes aleatorias

(8) despejar, concepto de prior y posterior

dar noción de costo (relación con BIC, AIK)

(9) kernel embedding

(10) Revisar bien el paper de clasificación

(11) Revisar bien la noción de tomar transporte triangular

Acta de Reunión