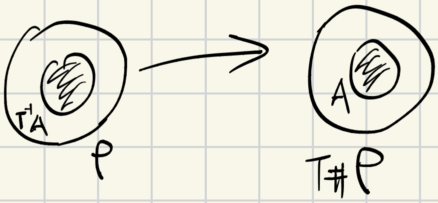
数学基础

邻域： $O_x^\varepsilon = \{ y : d(x,y) < \varepsilon \}$



开集： U ， $\forall x \in U$ ， $\exists \varepsilon$ ， $O_x^\varepsilon \subset U$

闭集： $U \Leftrightarrow X \setminus U$ 是开集

推前算子： $\int_{y \in A} T\#\rho(y)\,dy = \int_{T(x) \in A} \rho(x)\,dx$

变量替换

$T\#\rho(T(x))\,|T(x)| = \rho(x)$

$\rho(x)dx = \tilde{\rho}(T(x))dT(x)$

搬运沙子

概率测度空间的度量

$\rho, \rho' \in C^\infty$

① 全变差距离

$d_{TV}(\rho, \rho') = 0 \Leftrightarrow \|\rho - \rho'\|_{L_1} = 0 \Leftrightarrow \rho - \rho' = 0$

$d_{TV}(\rho, \rho') = d_{TV}(\rho', \rho')$

$$
\begin{aligned}
d_{TV}(\rho_1, \rho_2) &= \frac{1}{2}\int |\rho_1 - \rho_2|\,d\theta \\
&= \frac{1}{2}\int |\rho_1 - \rho_3 + \rho_3 - \rho_2|\,d\theta \\
&\le \frac{1}{2}\int |\rho_1 - \rho_3|\,d\theta + \frac{1}{2}\int |\rho_3 - \rho_2|\,d\theta
\end{aligned}
$$

② Hellinger 距离

$$
\begin{aligned}
d_H(\rho_1, \rho_2)^2 &= \frac{1}{2}\int |\sqrt{\rho_1} - \sqrt{\rho_2}|^2\,d\theta \\
&= \frac{1}{2}\int |\sqrt{\rho_1} - \sqrt{\rho_3} + \sqrt{\rho_3} - \sqrt{\rho_2}|^2\,d\theta \\
&= \frac{1}{2}\int |\sqrt{\rho_1} - \sqrt{\rho_3}|^2\,d\theta + \frac{1}{2}\int |\sqrt{\rho_3} - \sqrt{\rho_2}|^2\,d\theta \\
&\quad + \int |\sqrt{\rho_1} - \sqrt{\rho_3}|\,|\sqrt{\rho_3} - \sqrt{\rho_2}|\,d\theta
\end{aligned}
$$

$$= d_H(\rho_1, \rho_3)^2 + d_H(\rho_3, \rho_2)^2$$

$$+ \int \left|\sqrt{\rho_1} - \sqrt{\rho_3}\right| \left|\sqrt{\rho_3} - \sqrt{\rho_2}\right| d\theta$$

下证：

$$\int \left|\sqrt{\rho_1} - \sqrt{\rho_3}\right| \left|\sqrt{\rho_3} - \sqrt{\rho_2}\right| d\theta \le 2\, d_H(\rho_1\ \rho_3)\, d_H(\rho_3\ \rho_2)$$

使用 Cauchy 不等式

$$\int |fg| \le \sqrt{\int f^2 d\theta}\ \sqrt{\int g^2 d\theta}$$

良定性：

$$d_{TV}(\rho, \rho') = \frac{1}{2}\int |\rho - \rho'|\, d\theta \le \frac{1}{2}\Big(\underbrace{\int |\rho|\, d\theta + \int |\rho'|\, d\theta}_{=1}$$

$$d_H(\rho, \rho') = \left(\frac{1}{2}\int \left|\sqrt{\rho} - \sqrt{\rho'}\right|^2 d\theta\right)^{\frac{1}{2}}$$

$$= \left(\frac{1}{2}\int \rho + \rho' - 2\sqrt{\rho\rho'}\, d\theta\right)^{\frac{1}{2}}$$

$$\le 1$$

为什么不用 L_2 距离？

等价性：

$$\frac{1}{\sqrt{2}} d_{TV}(\rho, \rho') = \frac{1}{2\sqrt{2}} \int \left(\sqrt{\rho} + \sqrt{\rho'}\right) \left|\sqrt{\rho} - \sqrt{\rho'}\right| d\theta$$

$$\le \frac{1}{2\sqrt{2}} \sqrt{\int (\sqrt{\rho}+\sqrt{\rho'})^2 d\theta \int (\sqrt{\rho}-\sqrt{\rho'})^2 d\theta}$$

$$\le \frac{1}{2} \sqrt{4}\ d_H(\rho, \rho')$$

$$d_H(\rho,\rho')^2 = \frac{1}{2} \int |\sqrt{\rho} - \sqrt{\rho'}|^2 d\theta$$

$$d_{TV}(\rho,\rho') = \frac{1}{2} \int |\rho - \rho'| \, d\theta$$

由于 $|\sqrt{\rho} - \sqrt{\rho'}|^2 \le |\sqrt{\rho} - \sqrt{\rho'}| \, |\sqrt{\rho} + \sqrt{\rho'}|$

$$= |\rho - \rho'|$$

其它估计

首先证明

$$\frac{1}{2} |E_\rho f - E_{\rho'} f| \le d_{TV}(\rho, \rho')$$

$$\Longleftrightarrow \left|\int f(\rho-\rho')\, d\theta\right| \le \int |f| \, |\rho - \rho'| \, d\theta \qquad |f|_\infty \le 1$$

$$\le \int |\rho - \rho'| \, d\theta$$

再证明

$$\sup_{|f|_\infty \le 1} \frac{1}{2} |E_\rho f - E_{\rho'} f| \ge d_{TV}(\rho, \rho')$$

取 $f = \mathrm{sign}(\rho - \rho')$

$$\left| E_\rho f - E_{\rho'} f \right| = \left| \int f(\rho - \rho') d\theta \right|$$

$$\leq \int |f| \, |\rho - \rho'| \, d\theta$$

$$\leq 2 |f|_\infty \, d_{TV}(\rho, \rho')$$

$$\left| E_\rho f - E_{\rho'} f \right| = \left| \int f (\sqrt{\rho} - \sqrt{\rho'})(\sqrt{\rho} + \sqrt{\rho'}) \, d\theta \right|$$

$$\leq \left(\int f^2 (\sqrt{\rho} + \sqrt{\rho'})^2 d\theta \int (\sqrt{\rho} - \sqrt{\rho'})^2 d\theta \right)^{\frac{1}{2}}$$

$$\leq \left(\int f^2 2(\rho + \rho') \, d\theta \;\, 2 \right)^{\frac{1}{2}} d_H(\rho, \rho')$$

最优传输问题

Kantorovich 问题，有 $\gamma(\theta_1,\theta_2)$ 的沙子从 θ_1 运到 θ_2。

Kantorovich 对偶，引入拉格朗日乘子，f，g(离散)

$$\Phi(f,g) = \inf_{\gamma_{ij}\geq 0}\left\{ \sum_{ij} c_{ij}\gamma_{ij} + \sum_i f_i\left(a_i - \sum_j \gamma_{ij}\right) + \sum_j g_j\left(b_j - \sum_i \gamma_{ij}\right)\right\}$$

$$= \inf_{\gamma_{ij}\geq 0}\left\{ \sum_i f_i a_i + \sum_j g_j b_j + \sum_{ij}\gamma_{ij}\left(c_{ij} - f_i - g_j\right)\right\}$$

① Kantorovich $= \inf\limits_{\gamma_{ij}\geq 0 \ \sum\gamma_{ij}=a_i \ \sum\gamma_{ij}=b_j} \{ \quad \} \geq \Phi(f,g)$

$$\text{Kantorovich} \geq \sup_{f,g}\Phi(f,g) = \sup_{f_i+g_i\leq c_{ij}}\Phi(f,g)$$

$$= \sup_{f_i+g_i\leq c_{ij}}\left\{\sum_i f_i a_i + \sum_j g_j b_j\right\}$$

② 线性归化

$$\min\{c^T x : Ax=b \quad x\geq 0\}$$

$$\max\{b^T y : A^T y \leq c\}$$

下面证明 Wasserstein-p 距离是距离

假设我们有 ρ_A ρ_B ρ_C

$$W_p(\rho_A, \rho_C) \le \underset{\gamma_{AB}^*}{W_p(\rho_A, \rho_B)} + \underset{\gamma_{BC}^*}{W_p(\rho_B, \rho_C)}$$

那么有 $\gamma(x_A\ x_B\ x_C)$ 的边缘分布为 γ_{AB}^* γ_{BC}^*

$x_B \sim \rho_B$ $\quad x_A \sim \rho(\cdot \mid x_B)$ 根据 γ_{AB}^*

再生成 $\quad x_C \sim \rho(\cdot \mid x_B)$ 根据 γ_{BC}^*

$$W_p(\rho_A, \rho_C) \le \left(\iint \|x_A - x_C\|_2^p \, \gamma(x_A\ x_B\ x_C)\, dx_A\, dx_B\, dx_C\right)^{\frac{1}{p}}$$

$$\le \left(\int \{\|x_A - x_B\|_2 + \|x_B - x_C\|_2\}^p \gamma(x_A\ x_B\ x_C)\right)^{\frac{1}{p}}$$

$$\le W_p(\rho_A, \rho_B) + W_p(\rho_B, \rho_C)$$

这里用了 $p \ge 1$

Minkowski 不等式

Holder 不等式 $\quad \left(\int (\|f\|_2 + \|g\|_2)^p \gamma \, d\theta\right)^{\frac{1}{p}}$

$$\le \left(\int \|f\|_2^p \gamma \, d\theta\right)^{\frac{1}{p}} + \left(\int \|g\|_2^p \gamma \, d\theta\right)^{\frac{1}{p}}$$

思考: $\frac{1}{2}$ $\frac{1}{2}$ → $\frac{1}{2}$ $\frac{1}{2}$

$-1 \quad 0 \quad 1 \quad 2$

$W_1 = 2$

$W_2 = \sqrt{2^2 \cdot \left(\frac{1}{2} + \frac{1}{2}\right)} = 2$

对 $\quad c = \|\theta_1 - \theta_2\|_2$ ，下界的证明

$$\sup_h \int (\rho_A(\theta) - \rho_B(\theta))\, h(\theta)\, d\theta$$

$$= \sup_h \int \gamma(\theta_1, \theta_2)\left(h(\theta_2) - h(\theta_1)\right) d\theta_1 d\theta_2$$

$$\le \sup_h \int \gamma(\theta_1, \theta_2) \, |h(\theta_2) - h(\theta_1)| \, d\theta_1 d\theta_2$$

$$\le \sup_h \int \gamma(\theta_1, \theta_2) \, \|\theta_1 - \theta_2\|_2 \, d\theta_1 d\theta_2$$

$$\le W_1(\rho_A \; \rho_B)$$

上界的证明，目标

$$\sup_{f(\theta_1)+g(\theta_2) \le \|\theta_1-\theta_2\|_2} E_{\rho_A} f + E_{\rho_B} g$$

$$\le \sup_{|h(\theta_1) - h(\theta_2)| \le \|\theta_1 - \theta_2\|_2} E_{\rho_A} h - E_{\rho_B} h$$

构造

$$k(\theta) = \inf_u \left[\|\theta - u\|_2 - g(u) \right]$$ 由于 $f(\theta_1) + g(\theta_2) \le \|\theta_1 - \theta_2\|_2$

$$f(\theta_1) \le \inf_{\theta_2} \|\theta_1 - \theta_2\|_2 - g(\theta_2) = k(\theta_1)$$

$$k(\theta_2) = \inf_{\theta_1} \|\theta_2 - \theta_1\|_2 - g(\theta_1)$$

$$\le \|\theta_2 - \theta_2\|_2 - g(\theta_2) = -g(\theta_2)$$

我们有

$$\sup_{f(\theta_1)+g(\theta_2)\le \|\theta_1-\theta_2\|_2} E_{\rho_A} f + E_{\rho_B} g$$

$$\le E_{\rho_A} k - E_{\rho_B} k$$

下证 $|k(\theta_1) - k(\theta_2)| \le \|\theta_1 - \theta_2\|_2$

$$k(\theta_1) = \inf_u \|\theta_1 - u\|_2 - g(u)$$

$$\le \inf_u \|\theta_1 - u\|_2 - g(u)$$

$$\le \inf_u \|\theta_1 - \theta_2\|_2 + \|\theta_2 - u\|_2 - g(u)$$

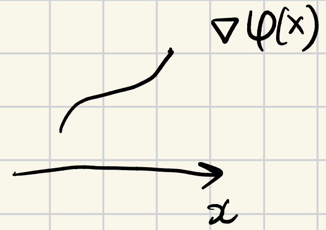
$$\le \|\theta_1 - \theta_2\|_2 + k(\theta_2)$$

因此 $k(\theta_1) - k(\theta_2) \le \|\theta_1 - \theta_2\|_2$

再由对称性可证。

$$(I \ \nabla\varphi): x \to (x, \nabla\varphi(x))$$

$$(I \ \nabla\varphi): P_A(x) \to \gamma(x, y)$$



$T: x \to y$

$T\# P_A \to P_B$

$(I, T)_\# P_A \to \gamma$

Wasserstein 2 距离

测地线

$$T_0 : (\theta_1\ \theta_2) \to \theta_1, \quad T_0 \# \gamma^* (A) = \gamma^* (T_0^{-1}(A))$$

$$= \gamma^* (\theta_1 \in A) = \rho_A$$

同理 $T_1 \# \gamma^* = \rho_B$

下证 $W_2(\rho_s, \rho_t) \leq (t-s) W_2(\rho_A, \rho_B)\ (s \leq t)$

定义 $\gamma^*_{s,t} = (T_s, T_t) \# \gamma^*$

$$(\theta_1, \theta_2) \to ((1-s)\theta_1 + s\theta_2, (1-t)\theta_1 + t\theta_2)$$

验证是耦合

$$\iint_{(\theta_1\ \theta_2) \in (A, R^d)} \gamma^*_{s,t}(\theta_1\ \theta_2)\, d\theta_2 = \int_{(\theta_1 \theta_2) \in (A, R^d)} \gamma^*_{s,t}(\theta_1, \theta_2)\, d\theta_1 d\theta_2$$

$$= \int_{\substack{[(1-s)x_1 + s x_2, \\ (1-t)x_1 + t x_2] \in (A, R^d)}} \gamma^*(x_1, x_2)\, dx_1 dx_2$$

$$= \int_{\substack{(1-s)x_1 + s x_2 \in A \\ (1-t)x_1 + t x_2 \in R^d}} \gamma^*(x_1, x_2)\, dx_1 dx_2$$

$$\int_{\theta_1 \in A} \rho_s(\theta_1)\, d\theta_1 = \int_{(1-s)\theta_1 + s\theta_2 \in A} \gamma^*(\theta_1, \theta_2)\, d\theta_2 d\theta_1$$

$$W_2(\rho_s, \rho_t)^2 \le \int \|\theta_1 - \theta_2\|^2 \gamma^*_{st}(\theta_1, \theta_2)\, d\theta_1 d\theta_2$$

换元

$$= \int \|T_s(\theta_1, \theta_2) - T_t(\theta_1, \theta_2)\|^2 \gamma^*(\theta_1, \theta_2)\, d\theta_1 d\theta_2$$

$$= (s-t)^2 \int \|\theta_1 - \theta_2\|^2 \gamma^*(\theta_1, \theta_2)\, d\theta_1 d\theta_2$$

$$= (s-t)^2 W_2(\rho_A, \rho_B)$$

因此

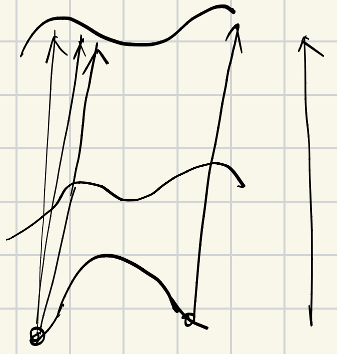
$$W_2(\rho_A, \rho_B) \le W_2(\rho_A, \rho_s) + W_2(\rho_s, \rho_t) + W_2(\rho_t, \rho_B)$$

$$\le s W_2(\rho_A, \rho_B) + (t-s) W_2(\rho_A, \rho_B) + (1-t) W_2(\rho_A, \rho_B)$$

$$= W_2(\rho_A, \rho_B)$$

所以全部取等

这对 W_p 距离也成立。



t

动力学观点

两点之间的距离

$X(s) = (1-s)X_0 + sX_1$ 满足

$$\inf_X \int_0^1 X'(s)^2 \, ds$$ （能量最小）

每一项放缩

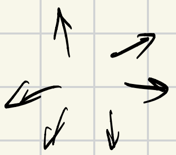
$$\int_0^1 X'(s)^2 ds \int_0^1 1 \, ds \geq \sum_i \left(\int X_i'(s) ds\right)^2$$

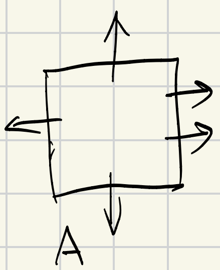
$$= \|X_1 - X_0\|_2^2$$

$$\int_0^1 (X'(s))^2 ds \geq \|X_1 - X_0\|_2^2$$

两个概率密度之间，有速度场 v_t

粒子随着 v_t 演化，ρ_t 也随之演化





$$\frac{\partial}{\partial t}\int_A \rho_t(\theta)\, d\theta = -\int_{\partial\Omega} \rho v \cdot n \, d\theta$$

$$= -\int_\Omega \nabla \rho_t v_t \, d\theta$$

$$\Rightarrow \partial_t \rho_t + \nabla_\theta(\rho_t v_t) = 0$$

定义 $T_t: R^d \to R^d$ 满足

$$\partial_t T_t(\theta) = v_t(T_t(\theta)), \quad T_0(\theta) = \theta$$

那么 $\rho_t = T_t \# \rho_0$，在时刻 t，粒子

从 $\theta \to T_t(\theta)$

因为粒子 $T_t(\theta)$ 下一时刻 $T_{t+\Delta t}(\theta)$

速度为 $v_t(T_t(\theta))$

$$\rho_0(\theta)\,d\theta = \rho_t(T_t(\theta))\,dT_t(\theta)$$

给定 v_t，在时间 0 到 1，把 $\rho_0 = \rho_A$ 演化

到 $\rho_1 = \rho_B$，定义能量

$$A(\rho, v) = \int_0^1 \int \|v_t\|_2^2\, \rho_t\, d\theta\, dt$$

① 我们有

$$= \int_0^1 \int \|v_t(T_t(\theta))\|_2^2\, \rho_t(T_t(\theta))\, dT_t(\theta)\, dt$$

$$= \int_0^1 \int \left\| \frac{\partial}{\partial t} T_t(\theta) \right\|_2^2 \rho_0(\theta)\, d\theta\, dt$$

由于 $\int_0^1 \left\| \frac{\partial}{\partial t} T_t(\theta) \right\|_2^2 dt \quad \int_0^1 dt$

每一项放缩

$$\geq \left(\int_0^1 \frac{\partial}{\partial t} T_t(\theta)\, dt \right)^2 = \| T_1(\theta) - T_0(\theta) \|^2$$

$$\geq \int \| T_1(\theta) - T_0(\theta) \|_2^2\, \rho_0(\theta)\, d\theta$$

$$\geq W_2^2(\rho_A, \rho_B)$$

② 另一方面，如果有最优映射 T，定义

$$T_t(\theta) = (1-t)\theta + tT(\theta) \qquad (T = \nabla\varphi)$$

定义 $v_t = \frac{d}{dt}T_t(\theta) \circ T_t^{-1}(\theta)$

$$= (T - Id) \circ T_t^{-1}$$

我们有 $\frac{d}{dt}T_t(\theta) = v_t(T_t(\theta))$，$\rho_t, v_t$

满足连续性方程，且

$$A(\rho, v) = \int_0^1 \int \left\| \frac{\partial}{\partial t} T_t(\theta) \right\|_2^2 \rho_0(\theta)\, d\theta\, dt$$

$$= \int \| T(\theta) - \theta \|_2^2 \, \rho_0(\theta)\, d\theta$$

$$= W_2^2(\rho_A, \rho_B)$$

KL 散度 $\rho^* = \frac{1}{Z} e^{-\Phi_R(\theta)}$

$$KL[\rho \| \rho^*] = \int \rho \log \frac{\rho}{e^{-\Phi_R}} + \rho \log Z \, d\theta$$

$$= \int \rho \log \rho + \rho \log \Phi_R \, d\theta + \log Z$$

KL 散度的强凸性

$$(1-t)KL[\rho_0\|\rho^*] + tKL[\rho_1\|\rho^*] - KL[(1-t)\rho_0+t\rho_1\|\rho^*]$$

$$\geq \frac{(1-t)t}{2}\|\rho_0-\rho_1\|_1^2$$

$$f(x) = x\ln x$$

$$KL[\rho\|\rho^*] = \int \rho^* f(\frac{\rho}{\rho^*})\,d\theta$$

$$f'(x) = \ln x + 1$$

$$f''(x) = \frac{1}{x}$$

定义

$$x_t = \frac{\rho_t}{\rho_*} = \frac{(1-t)\rho_0 + t\rho_1}{\rho_*}$$

$$a_i(\eta) = f((1-\eta)x_t + \eta x_i)$$

$$a_i(\eta) = a_i(0) + a_i'(0)\eta + \int_0^\eta a_i''(\tau)(\eta-\tau)\,d\tau$$

$\eta = 1$

$\Longrightarrow$

$$f(x_i) = f(x_t) + f'(x_t)(x_i-x_t) + (x_i-x_t)^2\int_0^1 f''((1-\tau)x_t+\tau x_i)(1-\tau)\,d\tau$$

$$(1-t)f(x_0) + tf(x_1)$$

$$= f(x_t) + (1-t)(x_0-x_t)^2\int_0^1 f''((1-\tau)x_t+\tau x_0)(1-\tau)\,d\tau$$

$$+ t(x_1-x_t)^2\int_0^1 f''((1-\tau)x_t+\tau x_1)(1-\tau)\,d\tau$$

$$= f(x_t) + (1-t)t \underbrace{(x_0-x_1)^2 \int_0^1 \left(\frac{t}{(1-\tau)x_t + \tau x_0} + \frac{1-t}{(1-\tau)x_t+\tau x_1} \right)(1-\tau)\,d}$$

下证

$$\int (\rho_0-\rho_1)^2 \int_0^1 \left(\frac{t}{(1-\tau)\rho_t+\tau\rho_0} + \frac{1-t}{(1-\tau)\rho_t+\tau\rho_1} \right)(1-\tau)\,d$$

$$= \int_0^1 (1-\tau)\,d\tau \int \frac{t(\rho_0-\rho_1)^2}{(1-\tau)\rho_t+\tau\rho_0} + \frac{(1-t)(\rho_0-\rho_1)^2}{(1-\tau)\rho_t+\tau\rho_1}\,d\theta$$

$$\geq \int_0^1 (1-\tau)\,d\tau \Big[t \left(\int |\rho_0-\rho_1|\,d\theta\right)^2 + (1-t)\left(\int |\rho_0-\rho_1|\,d\theta\right)^2 \Big]$$

$$= \frac{1}{2}\left(\int |\rho_0-\rho_1|\,d\theta\right)^2$$

随机过程：

$$E\, dB_t = E\, B_{t+dt} - B_t = 0$$

$$E\, dB_t^2 = E\,(B_{t+dt} - B_t)(B_{t+dt} - B_t) \approx dt$$

Itô 公式

$$dX_t = \partial_t f(t, \theta_t)dt + \nabla_\theta f(t, \theta_t)\, d\theta_t$$
$$+ \frac{1}{2} \nabla_\theta \nabla_\theta f(t, \theta_t)\, d\theta_t \cdot d\theta_t$$
$$= \partial_t f(t, \theta_t)dt + \nabla_\theta f(t, \theta_t)\left(b_t dt + \sigma_t dB_t\right)$$
$$+ \frac{1}{2}(\sigma_t dB_t)^T \nabla_\theta^2 f(t, \theta_t)\, (\sigma_t dB_t)$$
$$= \partial_t f(t, \theta_t)\, dt + \nabla_\theta f(t, \theta_t)\, b_t \cdot dt + \nabla_\theta f(t, \theta_t)\, \sigma_t dB$$
$$+ \frac{1}{2}\, dB_t^T\, \sigma_t^T\, \nabla_\theta^2 f(t, \theta_t)\, \sigma_t\, dB_t$$

$$\frac{1}{2} \nabla_\theta^2 f(t, \theta_t)\, \underline{(\sigma_t dB_t dB_t^T \sigma_t^T)}$$

$$dB_t \cdot dB_t^T \approx dt\, I$$

Fokker Planck 方程

$$d f(\theta_t) = \nabla_\theta f(\theta_t) b_t dt + \frac{1}{2} \nabla_\theta^2 f(\theta_t) : \sigma_t \sigma_t^T dt$$
$$+ \nabla_\theta f(\theta_t)^T \sigma_t dB_t$$

$$E f(\theta_t) = \int f(\theta) \rho_t(\theta) d\theta$$
$$= \int f(\theta_t) \rho_0(\theta) d\theta \quad \theta_t = \theta_t(\theta)$$

$$\partial_t E f(\theta_t) = E\left[\nabla_\theta f(\theta_t) b_t + \frac{1}{2} \nabla_\theta^2 f(\theta_t) : \sigma_t \sigma_t^T\right]$$
$$= \int \rho_t \left(\nabla_\theta f(\theta) b_t + \frac{1}{2} \nabla_\theta^2 f(\theta) : \sigma_t \sigma_t^T\right) d\theta$$
$$= -\int f(\theta) \nabla_\theta (\rho_t b_t) d\theta + \int \partial_{ij} f(\theta) \rho_t(\theta) D_{ij} d\theta$$
$$= -\int f(\theta) \nabla_\theta (\rho_t b_t) d\theta + \int f(\theta) \partial_{ij} (\rho_t(\theta) D_{ij}) d\theta$$

$$\Rightarrow \frac{\partial}{\partial t} \rho_t(\theta) = -\nabla_\theta (\rho_t b_t) + \sum \partial_{ij} (\rho_t D_{ij})$$

带约束问题

$$\min f(x)$$
$$g(x)=0$$
$$h(x)\le 0$$

P

拉格朗日乘子

Q

$$L(x,\lambda_g\ \lambda_h) = f(x)+\lambda_g\, g(x)+\lambda_h\, h(x)$$

$$\max_{\lambda_g\,\lambda_h}\ \min_{x}\ L(x,\lambda_g\ \lambda_h)$$

$$\lambda_h \ge 0$$

① 对于P的极小 x^*

$$f(x^*) \ge \min_{\substack{g(x)=0\\ h(x)\le 0\\ \lambda_h\ge 0}} L(x,\lambda_g\ \lambda_h)$$

因为 $f(x^*) \ge \underset{\lambda_h\ge 0}{L(x^*,\lambda_g,\lambda_h)}$

② 如果 P 是凸问题，那么两个问题极值相等

下面有KKT条件

$$\partial_x L(x^*, \lambda_g^*\ \lambda_h^*) = 0$$

$$g(x^*) = 0$$

$$h(x^*) \ge 0 \qquad \lambda_h^* h(x^*) = 0 \qquad \lambda_h^* \ge 0$$

KKT是充分条件 $\Rightarrow$ P

KKT是必要条件(如果P是凸) $\Leftarrow$ P

因此 x_2^* 满足问题1的条件，

$L_1 \leq L_2$

② 接下来我们考虑问题2的解

满足 $\nabla_x f(x) - \lambda_g \nabla_x g(x) - \lambda_h \nabla_x h(x) = 0$

$g(x) = 0$

$h(x) \geq 0$

是一个鞍点