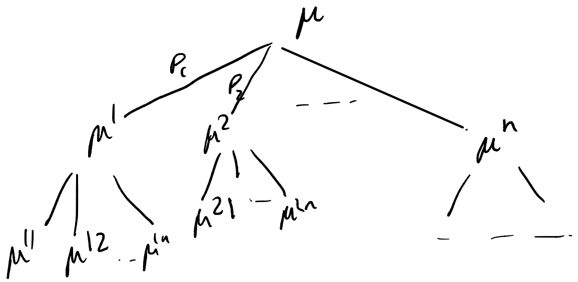
Stochastic Localization

We have been using a "localization" process for HDX:

$$\mu \text{ on } \binom{[n]}{k}: \quad \mu = p_1 \mu^1 + \cdots + p_n \mu^n$$

conditionals



Key: Every level we have a distribution over measures.

$$t=0 \quad \mathbb{E}[\mathrm{Var}_{\mu_1}[f]] = \chi^2(\nu \| \mu) - \chi^2(\nu D_{k\to 1} \| \mu D_{k \to 1})$$

$\vdots$

So if we think of a measure-valued random process:

$$\mu = \mu_0 \to \mu_1 \to \mu_2 \to \cdots \to \mu_k$$

Where to get μ_{t+1} we condition μ_t on a new element e w.p. $\propto$ marginals of link.

Then the process is a martingale:

$$\mathbb{E}[\mu_{t+1} \mid \text{past up to } t] = \mu_t$$

think of as vectors

HDX: $(\nu = f \cdot \mu)$

$$\mathbb{E}_t[\mathrm{Var}_{\mu_{t+1}}[f]] \geq \overset{1-\Omega(\frac{1}{k})}{\overset{\parallel}{\rho_t}} \cdot \mathrm{Var}_{\mu_t}[f]$$

$$\mathbb{E}_t[\mathrm{Ent}_{\mu_{t+1}}[f]] \geq \rho_t \cdot \mathrm{Ent}_{\mu_t}[f]$$

Local-to-Global:

$$\mathbb{E}[\mathrm{Var}_{\mu_\ell}(f)] \geq \rho_0 - \rho_{\ell-1} \mathrm{Var}_{\mu}[f]$$

$$\mathbb{E}[\mathrm{Var}_{\mu_\ell}(f)] = \chi^2(\nu \| \mu) - \chi^2(\nu D_{k\to\ell} \| \mu D_{k\to\ell})$$

$$\mathrm{Var}_{\mu}[f] = \chi^2(\nu \| \mu)$$

$$\mathbb{E}[\mathrm{Ent}_{\mu_\ell}[f]] \geq \rho_0 - \rho_{\ell-1} \mathrm{Ent}_{\mu}[f]$$

$$\mathbb{E}[\mathrm{Ent}_{\mu_\ell}[f]] = D_{KL}(\nu \| \mu) - D_{KL}(\nu D_{k\to\ell} \| \mu D_{k\to\ell})$$

$$\mathrm{Ent}_{\mu}[f] = D_{KL}(\nu \| \mu)$$

[Chen-Eldan] studied "localizatio schemes" more generally. The equiv. notions of HDX are called conservation of variance & conservation of entropy.

Stochastic localization is a different scheme that works in continuous time:

$$\{\mu_t\}_{t \in \mathbb{R}_{\geq 0}}$$

(For simplicity think of μ_t on $\{\pm 1\}^n$, but it works for μ_t on $\mathbb{R}^n$)

The original def is to multiply a linear:

$$d\mu_t(x) = \mu_t(x) \langle x - \mathrm{mean}(\mu_t), \underbrace{dW_t}_{\text{this is a random vector}} \rangle$$

Alternative:

$$\mu_{t+dt}(x) = \mu_t(x) \left(1 + \langle x - \mathrm{mean}(\mu_t), \underbrace{dW_t}_{\text{think of } \mathcal{N}(0, dt I)} \rangle\right)$$

Note that μ_{t+dt} is still a distribution

$$\sum_x \mu_{t+dt}(x) = \sum_x \mu_t(x) + \langle dW_t, \overset{\text{mean}}{\cancel{\sum_x \mu_t(x) x}} - \overset{\text{mean}}{\cancel{\sum_x \mu_t(x) \text{mean}}} \rangle$$

So as long as dW_t is infinitesimally small, so that no $\mu_t(x)$ becomes < 0,

μ_t will be a distribution at all times.

What about being a martingale?

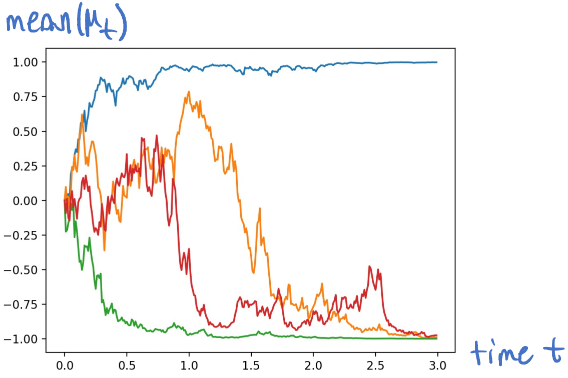
$\forall x$

$$\mathbb{E}_t[d\mu_t(x)] = \mu_t(x) \cdot \langle x - \text{mean}(\mu_t), \underbrace{\mathbb{E}_t[dW_t]}_{\text{enough to be 0}} \rangle$$

lem: As long as W_t is a martingale, that is $\mathbb{E}_t[dW_t] = 0$, the process μ_t is a dist-valued martingale. ☺

Convenient Choice: Take W_t to be Brownian motion:

$$dW_t = \mathcal{N}(0, dt I)$$



Four runs for μ unif on $\{\pm 1\}$

Note that we do not have to fix W_t in advance. In fact we can let

$$dW_t = \mathcal{N}(\underset{\text{necessary for martingale}}{0}, \overset{\text{decide based on } \mu_t}{\Sigma_t dt})$$

Can we write a closed-form expression for μ_t?

Claim: We always have

$$\mu_t(x) \propto \mu_0(x)\, e^{\underbrace{\frac{1}{2} x^T A_t x + h_t^T x}_{\text{exp of quadratic}}}$$

This follows from Itō's lemma:

Quick aside:

Suppose we have a cont. time process X_t that:

$$\underset{\text{Such a thing is called Itō process}}{dX_t} = m_t dt + \sigma_t \underbrace{dB_t}_{\text{standard Brownian motion}}$$

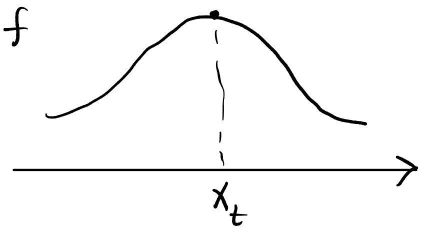
Then what is $Y_t = f(X_t)$ for some "nice function" f?

Chain rule is wrong:

$$dY_t \neq f'(X_t)(m_t dt + \sigma_t dB_t)$$

Example.

Y_{t+dt} is on average smaller!



Itō's lemma: We have

$$dY_t = f'(X_t)(m_t dt + \sigma_t dB_t) + \overbrace{\frac{1}{2} f''(X_t)\, \sigma_t^2 dt}^{\text{Itō term}}$$

This is because in Taylor series of f we have to go until second order.

$$f(X_t + m_t dt + \sigma_t dB_t) \simeq f(X_t) + f'(X_t)(m_t dt + \sigma_t dB_t) +$$
$$\frac{1}{2} f''(X_t)(m_t dt + \sigma_t dB_t)^2 + \cdots$$

We can drop everything of magnitude $\ll dt$

but $dB_t^2 \approx dt$.

We can use Itō to understand $\lg(\mu_t(x))$

in stochastic localization.

$$X_t = \mu_t(x)$$

$$dX_t = \mu_t(x)\left(\langle x - \mathrm{mean}(\mu_t), dW_t\rangle\right)$$

$$dW_t = \underline{C_t} \cdot \underbrace{dB_t}_{\text{n-dim Brownian}}$$

$\mathcal{N}(0, \Sigma_t dt)$

$Y_t = \lg \mu_t(x)$ then

$\lg'(z) = \frac{1}{z}$

$\lg''(z) = \frac{-1}{z^2}$

$$dY_t = \underbrace{\langle x - \mathrm{mean}(\mu_t), dW_t\rangle}_{\text{Chain rule part}} +$$

$$-\frac{dt}{\underbrace{2\cancel{X_t^2}}_{\frac{1}{2}f''(X_t)}} \cdot \cancel{X_t^2} \cdot \underbrace{(x - \mathrm{mean}(\mu_t))^T \boxed{C_t C_t^T} (x - \mathrm{mean}(\mu_t))}_{\text{variance}}$$

Σ_t

So $dY_t = -\frac{1}{2} x^T \Sigma_t x dt + \langle \text{something}, x\rangle + \text{something}$

This means that we always have

$$\mu_t(x) \propto \mu_0(x)\, e^{x^T A_t x + h_t^T x} \quad \text{and}$$

$$A_t = -\frac{1}{2}\int_0^t \Sigma_t \, dt$$

$I[i \to j]$
$\leq |J_{ij}|$
$J \to J + cI$

This is very convenient when μ_0 itself is exp(quadratic).

[Eldan-Koehler-Zeitouni]

Thm: Suppose $\mu(x) \propto \exp\left(\frac{1}{2} x^T J x + h^T x\right)$ on $\{\pm 1\}^n$. This is an Ising model. Then if $0 \preceq J \preceq (1-\delta) I$, Glauber dynamics on μ mixes rapidly ($t_{rel} = O_\delta(n)$).

- We also know MLSI [Anari-Koehler-Pham-Vuong]

Proof: Suppose we run stochastic localization. Then we have

$$\mu_t(x) \propto \exp\left(\frac{1}{2} x^T J_t x + h_t^T x\right)$$

$$J_t = J - \int_0^t \Sigma_t \, dt$$

We will make sure that $J_t \succeq 0$.

What we'd like to show is that for

$$\nu = f\mu$$

we have

$$\mathcal{E}(f,f) \geq \Omega_\delta\left(\frac{1}{n}\right) \mathrm{Var}_\mu[f]$$

dirichlet form of Glauber dynamics

$\mathbb{E}_{x\sim\mu}[f(x)^2] - \mathbb{E}_{x\sim\mu}[f(x)]^2$

Luckily the dirichlet form is a supermartingale.

$$\mathbb{E}_t\left[d\,\mathcal{E}_{\mu_t}(f,f)\right] \leqslant 0$$

$$\sum_{x\sim y} \underbrace{Q_{\mu_t}(x,y)}_{\frac{1}{n}\cdot\frac{\mu_t(x)\,\mu_t(y)}{\mu_t(x)+\mu_t(y)}} (f(x)-f(y))^2$$

on hypercube

this is a Harmonic mean

Exercise: Because $\mu_t(x)$, $\mu_t(y)$ are martingales their harmonic mean is supermartingale.

- So the Dirichlet form only decreases over time (on average)

- If we could prove $\mathrm{Var}_{\mu_t}[f]$ stays the same or increases, we could then pull back a Poincaré inequality for μ_t to a Poincaré inequality for μ_0.

- Note that $\mathrm{Var}_{\mu_t}[f] = \underbrace{\overbrace{\mathbb{E}_{x\sim\mu_t}[f(x)^2]}}_{\text{this stays on avg. the same}} - \underline{\mathbb{E}_{x\sim\mu_t}[f(x)]^2}$

- So it is enough to make sure that $\mathbb{E}[f(x)]^2$ stays the same on avg.

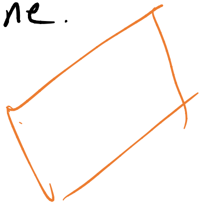
We will show with prob 1!

Let's keep $\mathbb{E}_{\mu_t}[f(x)]$ the same a.s.

$$d\mathbb{E}_{\mu_t}[f(x)] = \sum_x f(x)\, d\mu_t(x)$$

$$= \Big\langle \underbrace{\sum_x \mu_t(x) f(x)\, x - \sum_x \mu_t(x) \overset{f(x)\,\mathrm{mean}(\mu_t)}{\vee}}_{\text{some vector } v_t}, dW_t \Big\rangle$$

As long as we choose dW_t orthogonal to this vector we are fine.



Constraint: $\Sigma_t v_t = 0$

$$\mathrm{span}(\Sigma_t) \subseteq \underset{\substack{\downarrow \\ \text{to make sure } J_t \succeq 0}}{\mathrm{span}(J_t)}$$

This has a nonzero solution as long as rank $J_t > 1$.

So we run the process and reduce J_t until it becomes rank 1.

Thm: If J_t is rank 1 and $0 \preceq J_t \preceq (1-\delta)I$ then it has $\Omega_\delta(\frac{1}{n})$ Poincaré.

Proof: Follows from Dobrushin++.

$J_t = uu^T$ $\quad \|u\|^2 \leqslant 1-\delta$

$u'_i = |u_i|$ $\quad \|u'\|^2 = \|u\|^2 < 1-\delta$

influence $\leqslant u'u'^T \preceq (1-\delta)I$