$$MC = (\Omega, \nu_0, K)$$

- Ω is the state space
- ν_0 is initial state space distribution
- K is transition matrix $K : \Omega \times \Omega \to \mathbb{R}$

Goal:

Design MC s.t. $\lim_{n \to \infty} \nu_0 K^n = \overset{\text{target distribution}}{\pi}$

Properties of K

Let $\pi \sim \pi(i)$ be probability (think vector)

- Stochastic: $\sum_{j=1}^{N} K_{ij} = 1 \iff K I = I$
- Global Balance (stationarity): $\pi K = \pi$
  - $\sum_i \pi(i) K_{ij} = \pi(j) \quad \forall j$
  - detailed balance: $\pi(i) K_{ij} = \overset{\text{reversibility}}{\pi(j) K_{ji}}$

$\boxed{* \text{ det. bal} \rightarrow \text{gl. bal}}$

$$\sum_i \pi(i) K_{ij} = \sum_i \pi(j) K_{ji} = \pi(j) \sum_i K_{ji} = \pi(j)$$

- **Irreducibility**: all states are connected

- **Aperiodic**: $\nexists\, d>1$ s.t. $P_{ii}^{n} = 0$ whenever $n = 0 \mod(d)$, $\forall i$
  - for an irreducible chain, a single state i w/ $P_{ii}^{k} \neq 0$ for some k implies aperiodicity.

- **Stationarity**: $\exists \pi$ s.t. $\pi K = \pi$

Side note: Perron-Frobenius Theorem

$$K^n = I\pi + O(n^{m_2-1}|\lambda_2|^n)$$

$$m_2 = \text{mult}(\lambda_2) \quad \text{where} \quad 1 = \lambda_1 > |\lambda_2| > \dots$$

Metropolis-Hastings

Input: ① target dist. π

② proposal distribution $Q(\overset{\text{current state}}{x}, \underbrace{y}_{\text{proposed state}})$

$\sim Q_x(y) \sim$ pdf on y

③ $$\alpha(x,y) = \min\left[1, \frac{Q(y,x)}{Q(x,y)} \frac{\pi(y)}{\pi(x)}\right]$$

Idea: det. balance → global balance ⟷ stationarity

Theorem: MH satisfies detailed balance

Proof:

$$K(x,y) = \begin{cases} Q(x,y)\alpha(x,y) & ; \ x \neq y \\ 1 - \sum_{x \neq y} Q(x,y)\alpha(x,y) & ; \ y = x \end{cases}$$

Note that if $\alpha(x,y) \neq 1$, then $\alpha(y,x) = 1$

Say $x \neq y$:

$$\pi(x) K(x,y) = \pi(x)\, Q(x,y)\, \overset{\min\left(1, \frac{Q(y,x)\pi(y)}{Q(x,y)\pi(x)}\right)}{\alpha(x,y)}$$

$$= \min\left(\pi(x) Q(x,y),\ Q(y,x)\pi(y)\right)$$

$$\pi(y) K(y,x) = \min\left(\pi(y) Q(y,x),\ Q(x,y)\pi(x)\right)$$

▨

Alternate MH:

Take Q symmetric ~ $Q(x,y) = Q(y,x)$

$$\Rightarrow \alpha = \min\left(1, \frac{\pi(y)}{\pi(x)}\right)$$

Note: $\alpha(x,y) = \dfrac{s(x,y)}{Q(x,y)\,\pi(x)}$ for $s(x,y) = s(y,x)$

is a general form of α.

Key Issues:

① $\forall x$, $\{y \mid Q(x,y) > 0\}$ is large

so that $K(x,y)$ is well connected

② $\forall x$, $\text{prob}(y) = Q(x,y)$ is far from

uniform - Well Informed.

Autocorrelation

$$C_{ij} = \langle f_i f_j \rangle - \langle f_i \rangle \langle f_j \rangle$$

where $f_i = f(x_i)$ when $x_i \sim MC$

$$E[\overline{f(x)}_n] = \frac{1}{n} \sum_{i=1}^{n} E[f(x_i)] = \frac{1}{n} \sum_{i} E[f(x)]$$

$$= E[f(x)]$$

$$Var[\overline{f(x)}_n] = E\left[\left(\overline{f(x)}_n - E(\overline{f(x)})\right)^2\right]$$

$$= E\left[\left(\overline{f(x)}_n - E(f(x))\right)^2\right]$$

$$= E\left[\left(\frac{1}{n}\sum_i f(x_i) - \overset{\frac{1}{n}\sum E(f(x))}{E(f(x))}\right)\left(\frac{1}{n}\sum_j f(x_j) - E(f(x))\right)\right]$$

$$= E\left[\left(\frac{1}{n}\sum_i \left(f(x_i) - E(f(x))\right)\right)\left(\frac{1}{n}\sum_j \left(f(x_j) - E(f(x))\right)\right)\right]$$

$$= \frac{1}{n^2} E\left[\sum_{ij} (f(x_i) - \mu_f)(f(x_j) - \mu_f)\right]$$

$$= \frac{1}{n^2} E\left[\sum_i (f(x_i) - \mu_f)^2 + \sum_{i \neq j} (f(x_i) - \mu_f)(f(x_j) - \mu_f)\right]$$

$$= \frac{1}{n^2}\left(\sum_i E\left[(f(x_i)-\mu_f)^2\right] + \sum_{i\neq j} E\left[(f(x_i)-\mu_f)(f(x_j)-\mu_f)\right]\right)$$

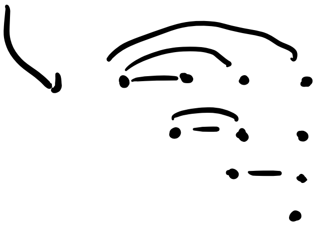
$$= \frac{1}{n^2}\left(\sum_i \overset{\overset{\mathrm{Var}(f(x))}{\sim}}{\mathrm{Var}}(f(x_i)) + 2\sum_{i=1}^{n}\sum_{j>i} \mathrm{Cov}(f(x_i), f(x_j))\right)$$

$$\mathrm{Var}(\overline{f(x)}_n) = \frac{1}{n}\mathrm{Var}(f(x)) + \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=i+1}\mathrm{Cov}(f(x_i), f(x_j))$$

$j = i+t$

$$\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathrm{Cov}(f_i, f_j) = \sum_{i=1}^{n}\sum_{t=1}^{n-i}\mathrm{Cov}(f_i, f_{i+t})$$

$$= \sum_{t=1}^{n-1}(n-t)\,C_t^f$$



$(n-1)C_1$

$(n-2)C_2$

$(n-3)C_3$

$\vdots$

$$\mathrm{Var}(\overline{f(x)}_n) = \frac{1}{n}\mathrm{Var}(f(x)) + \frac{2}{n^2}\sum_{t=1}^{n-1}(n-t)\,C_t^f$$

$$= \frac{1}{n}\mathrm{Var}(f(x))\left[1 + 2\sum_{t=1}^{n-1}\left(\frac{n-t}{n}\right)\rho_t^f\right]$$

where $\rho_t = \dfrac{C_t^f}{\sigma_f^{\,2}} \sim \dfrac{\mathrm{Cov}(x_i, x_j)}{\sigma_i\,\sigma_j}$

is autocorrelation function.

Sample autocovariance & autocorrelation

$$\gamma(t) = \frac{1}{n-t}\sum_{i=1}^{n-t}\left(f(x_i)-\overline{f(x)}_n\right)\left(f(x_{i+1})-\overline{f(x)}_n\right)$$

$$\gamma(0) = \frac{1}{n}\sum_{i=1}^{n}\left(f(x_i)-\overline{f(X)}_n\right)^2 = s^2$$

$\rho_s(t) = \dfrac{\gamma(t)}{\gamma(0)}$ is sample autocorrelation

thought: ① How are $\gamma(t)$, $\gamma(0)$, & $\rho_s(t)$ distributed?

② Are these estimators biased? MLE...? OLS?

③ $\gamma(t) = \frac{1}{n}\sum$ ... better for $t \sim n$?

$$¿\left(1 + 2\sum_{t=1}^{n-1}\frac{n-t}{n}\rho_s(t)\right) \approx 2\sum_{t=1}^{\infty} e^{-t/\tau_f} \approx 2\tau_f \,?$$

$$\Rightarrow \mathrm{Var}\left(\overline{f(X)}_n\right) \approx \frac{1}{n}\mathrm{Var}(f(X)) \cdot 2\tau_f$$

$$n_{eff} \approx n/2\tau_f$$