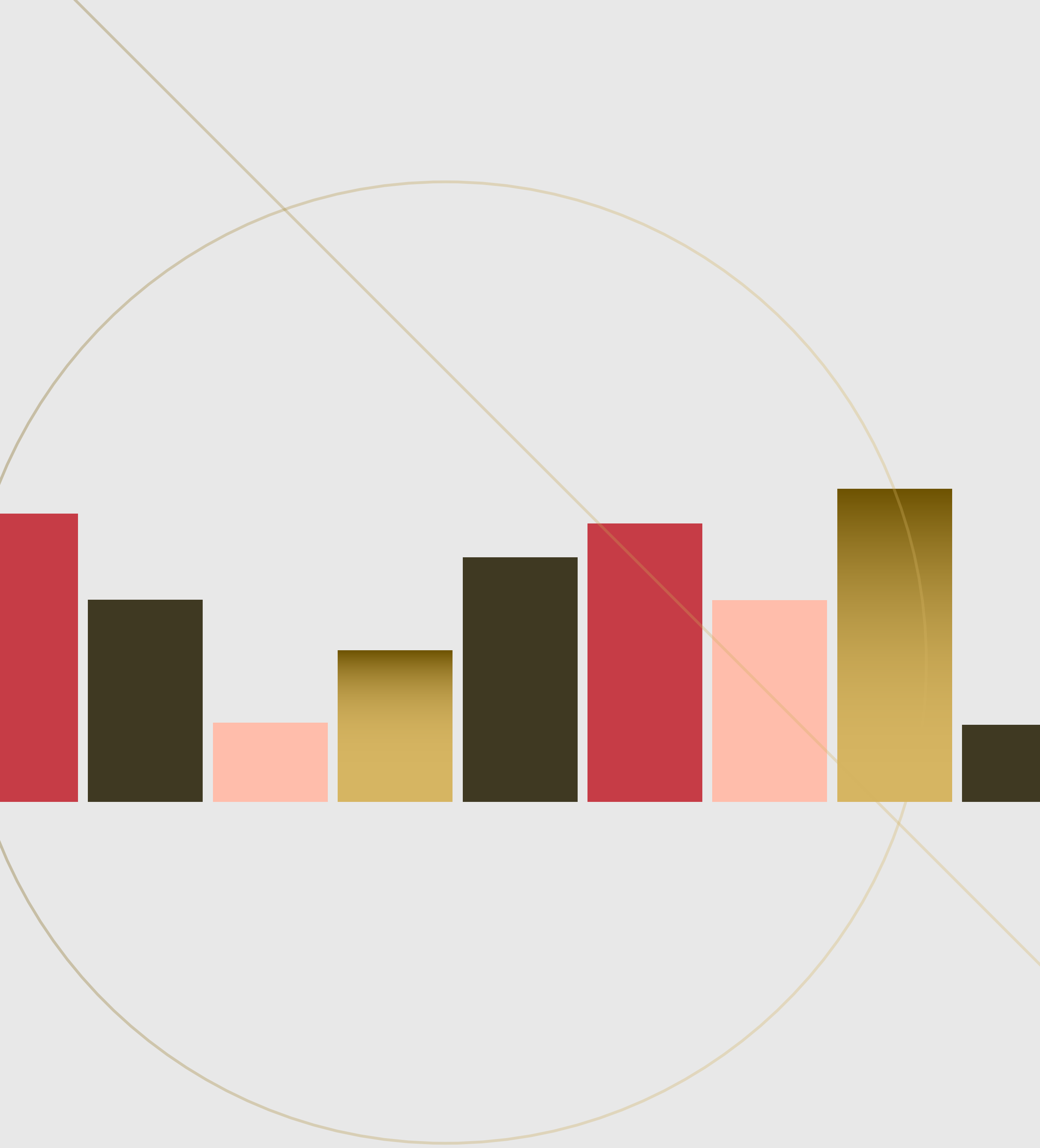


Markov Chains for exploring posterior distributions
- Tierney (1994)
The Annals of Statistics

π - posterior distribution
θ - parameters

Let π be an invariant distribution of time-homogen-eous Markov chain which is a sequence of random variables $\{X_n ; n \geq 0\}$

Transition kernel P is defined as

$$P(X_n, A) = P\{X_{n+1} \in A \mid X_0 .. X_n\}$$

which satisfies

$$\pi(A) = \int \pi(dx) P(x, A)$$

$P^n[X_0, A]$ = conditional distribution of X_n given X_0.
where P^n is n^{th} iterate of P.

If $\lim_{n \to \infty} P^n(x, A) = \pi(A)$, then π is invarient distribution to the Markov chain.

For π to be unique invarient distribution, the chain should be irreducible and aperiodic.

Gibbs Sampler

$X \sim \pi$

Let $Y = f(X)$

$$Q(y, A) = P\{X \in A \mid Y = y\}$$

$$\therefore \quad P(x, A) = Q(f(x), A)$$

Let's assume that $x = \{x_1, x_2 \cdots x_m)$

$$P(x, A) = \text{transition probability}$$
$$= (x_1 \mid f_1(x))(x_2 \mid f_2(x)) \cdots (x_m \mid f_m(x))$$
$$= \underbrace{(x_1 \mid x_2 \cdots x_m)}_{P_1} \underbrace{(x_2 \mid x_1, x_3 \cdots x_m)}_{P_2} \cdots \underbrace{(x_m \mid x_1 \cdots x_{m-1})}_{P_m}$$

$$P = P_1 P_2 \cdots P_m$$

where P_1 is a transitional probability which puts all it's weight to the form $(x_1, x_2^{(n)}, \cdots x_m^{(n)})$ where $x_2^{(n)}, \cdots x_m^{(n)}$ are fixed.

If we start with an initial distribution say
$x^{(0)} = (x_1^{(0)} \cdots x_m^{(0)})$

$$(x_1^{(1)} \cdots x_m^{(1)}) \sim (x_1 \mid x_2^{(0)} \cdots x_m^{(0)})(x_2 \mid x_1^{(1)}, x_3^{(0)} \cdots x_m^{(0)}) \cdots (x_m \mid x_1^{(1)} \cdots x_{m-1}^{(1)})$$
$$\sim (x_1 \mid f_1(x^{(0)}) \cdots (x_m \mid f_m(x^{(1)}))$$
$$\sim X \mid Y = f(X^{(0)})$$
$$\sim X \mid Y = y$$

Theoretical results

* $\mathcal{E}$ = σ algebra on some space E
* $P : E \times \mathcal{E} \longrightarrow [0,1]$ is a trasition kernal on $(E, \mathcal{E})$ s.t
  1) for any fixed $\underbrace{A}_{\text{set}} \in \underbrace{\mathcal{E}}_{\text{set of sets}}$, the function $P(\cdot, A)$ is measurable.
  2) for any $x \in E$, the function $P(x, \cdot)$ is a probability measure on $(E, \mathcal{E})$.

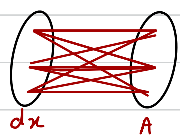
  $\therefore$ $(E, \mathcal{E}, P)$ is a probability space with measure P.

P_x = transition kernal P started at x

ν is probability

h is real valued $\mathcal{E}$ measurable function, then

$$(\nu P)(A) = \int P(x, A)\, \nu(dx)$$



$$(Ph)(x) = \int h(y)\, P(x, dy)$$

$$\nu h = \int h(y)\, \nu(dy)$$

h is called harmonic for P if $\underbrace{h = Ph}_{\text{similar to condition for invariant distribution}}$.

Total variation norm for bounded signed measure λ on $(E, \mathcal{E})$ is defined as

$$\| \lambda \| = \sup_{A \in \mathcal{E}} \lambda(A) - \inf_{A \in \mathcal{E}} \lambda(A)$$

Total variation distance b/w 2 such measures λ_1 and λ_2 are $\| \lambda_1 - \lambda_2 \|$

Irreducibility
It is defined w.r.t a σ-finite measure φ.
A transition kernal P on $(E, \mathcal{E})$ is φ-irreducible if $\varphi(E) > 0$ & for each $x \in E$ & each $A \in \mathcal{E}$ with $\varphi(A) > 0$ there exist an integer $n = n(x, A) \geq 1$ s.t $P^n(x, A) > 0$

We will assume $\varphi = \pi$ - target density

Periodicity
$\exists \; d \geq 2$ and a sequence $\{E_0, E_1, \ldots E_{d-1}\}$ of d non empty disjoint sets in $\mathcal{E}$ s.t, $\forall \; i = 0, \ldots, d-1$

$$P(x, E_j) = 1 \quad \text{for } j = i+1 \pmod d$$