Spectrum of reversible Markov chains, and mixing

Recall that the mixing time question is 'just' a question about the convergence of the 'power method'.

We wat to show that

$$\mu P^t \xrightarrow[t\uparrow\infty]{} \pi$$

$\forall$ probability distribution μ, and π is the 'top' eigenvector of P.

- The proof methods we have seen are more probabilistic.

<u>Simple setting</u> :- P is stochastic and symmetric.

- Symmetry implies that P has real eigenvalues with corresponding real <u>orthonormal</u> eigenvectors.

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots \geq \lambda_n$$
$$v^{(1)} \quad v^{(2)} \quad v^{(3)} \quad \cdots \quad v^{(n)}.$$

- Stochasticity implies $\lambda_1 = 1$, $|\lambda_i| \leq 1$. $1 = \lambda_1 \geq \lambda_n \geq -1$.

- In fact the distribution is a stationary distribution of P. In particular, all entries of $v^{(1)}$ are equal, so $v_i^{(1)} = \frac{1}{\sqrt{n}}$ $\forall 1 \leq i \leq n$.

We also get the representation

$$P = \sum_{i=1}^{n} \lambda_i v^{(i)} v^{(i)^T} = \frac{1}{n} J + \sum_{i=2}^{n} \lambda_i v^{(i)} v^{(i)^T}$$

J = $n \times n$ matrix with all entries equal to 1.

So for any probability distribution.

$$\mu^T P = \frac{1}{n} \mu^T J + \sum_{i=2}^{n} \lambda_i \langle \mu, v^{(i)} \rangle v^{(i)^T}$$

$$= \pi + \sum_{i=2}^{n} \lambda_i \langle \mu, v^{(i)} \rangle v^{(i)^T}$$

where $\pi = \left(\frac{1}{n}, \frac{1}{n}, \ldots \frac{1}{n}\right)$ is a stationary distribution of π.

$$\mu^T P - \pi = \sum_{i=2}^{n} \lambda_i \langle \mu, v^{(i)} \rangle v^{(i)^T}.$$

The same calulation repeated with P^k gives

$$\mu^T P^k - \pi = \sum_{i=2}^{n} \lambda_i^k \langle \mu, v^{(i)} \rangle v^{(i)^T}.$$

We wanted a bound on the TV-norm

$$\|\mu^T P^k - \pi\|_{TV} = \frac{1}{2} \|\mu^T P^k - \pi\|_1$$

$$\leq \frac{\sqrt{n}}{2} \|\mu^T P^k - \pi\|_2 \quad \text{when } P \text{ is } n \times n.$$

$$\|\mu^T P^k - \pi\|_2^2 = \left\| \sum_{i=2}^{n} \lambda_i^k \langle \mu, v^{(i)} \rangle \, v^{(i)T} \right\|_2^2 .$$

$$= \sum_{i=2}^{n} |\lambda_i|^{2k} |\langle \mu, v^{(i)} \rangle|^2$$

, because $(v^{(i)})_{i\geq 2}^{n}$ is a set of orthonormal vectors.

Define $\lambda := \max_{i\geq 2} \{ |\lambda_i| \}$.

$$\|\mu^T P^k - \pi\|_2^2 \leq \lambda^{2k} \sum_{i=2}^{n} |\langle \mu, v^{(i)} \rangle|^2$$

$$= \lambda^{2k} \left(\|\mu\|_2^2 - \langle \mu, v^{(1)} \rangle^2 \right)$$

$$\leq \lambda^{2k} \|\mu\|_2^2 \leq \lambda^{2k} \|\mu\|_1^2 \leq \lambda^{2k} .$$

$$\|\mu^T P^k - \pi\|_{TV} \leq \frac{1}{2}\|\mu^T P^k - \pi\|_1 \leq \frac{\sqrt{n}}{2}\|\mu^T P^k - \pi\|_2$$

$$\leq \frac{\sqrt{n}}{2}\lambda^k.$$

So if $\lambda < 1$, then,

$$\|\mu^T P^k - \pi\|_{TV} \leq \frac{1}{4} \text{ when } k \geq \frac{\log(2\sqrt{n})}{\log(1/\lambda)}.$$

Define $\gamma = 1 - \lambda$. "Spectral gap"

Then $\tau_{mix}\left(\frac{1}{4}\right) \leq \left\lceil \frac{\log(2\sqrt{n})}{\log\left(\frac{1}{1-\gamma}\right)} \right\rceil \approx O\left(\frac{1}{\gamma} \underline{\underline{\log n}}\right)$

So, Large Spectral gap $\Rightarrow$ Fast mixing.

Can be easily extended to the reversible case.

- There exist $\pi > 0$ s.t.

$$\pi(x) P(x,y) = \pi(y) P(y,x).$$

- In this case P is 'self-adjoint' with respect to the inner product defined by π:

$$\langle u, v \rangle_\pi := \sum_{x \in \Omega} \pi(x) u(x) v(x),$$ and has

a complete basis of eigenvectors that are orthonormal wrt the $\langle \cdot , \cdot \rangle_\pi$ inner product.

- Equivalent, one can argue as follows:- Reversibility Condition

$$\pi(x) P(x,y) = \pi(y) P(y,x) \quad \forall x, y$$

is the same as the matrix condition

$$D_\pi P = P^T D_\pi \iff D_\pi^{1/2} P D_\pi^{-1/2} = D_\pi^{-1/2} P^T D_\pi^{1/2}$$

when $\pi > 0$.

where D_π is the diagonal matrix for which $D_\pi(x,x) = \pi(x)$.

$$(D_\pi P)(x,y) = D_\pi(x,x) P(x,y) = \pi(x) P(x,y).$$

$$(P^T D_\pi)(x,y) = P^T(x,y) D_\pi(y,y) = P(y,x)\pi(y).$$

Let's consider $\underbrace{D_\pi^{1/2} P D_\pi^{-1/2}}_{Q} = \underbrace{D_\pi^{-1/2} P^T D_\pi^{1/2}}_{Q^T}$.

So, $Q = Q^T$ where $Q = D_\pi^{1/2} P D_\pi^{-1/2}$ when $\pi > 0$ and P is reversible with respect to π. So, Q is symmetric & has real eigenvalues.

- What are the eigenvalues of Q? $\left(Q = D_\pi^{1/2} P D_\pi^{-1/2}\right)$
  $\pi > 0$.
- Is Q stochastic?

$$Q(x,y) = \sqrt{\frac{\pi(x)}{\pi(y)}}\, P(x,y).$$

Q has the same set of eigenvalues as P, counted with multiplicity.

Aside:

$$Tv = \lambda v$$

(Suppose R, T, S are square matrices, where R is invertible and $T = RSR^{-1}$.)

$$\iff RSR^{-1}v = \lambda v$$

$$\iff S(R^{-1}v) = \lambda(R^{-1}v)$$

So, v is an eigenvector of T with eigenvalue v

iff

$R^{-1}v$ is an eigenvector of T with eigenvalue v.

So, $D_\pi^{1/2} v$ is an eigenvector of Q iff v is an eigenvector of P. They have the same eigenvalues.

Since the eigenvectors of Q can be chosen to be an orthonormal basis we get that P has a complete basis of eigenvectors $v_1^{(1)}, v^{(2)}, \ldots v^{(n)}$ s.t.

$$\left\langle v_1^{(i)} D_\pi v_1^{(j)} \right\rangle = \begin{cases} 1 & \text{when } i=j \\ 0 & \text{otherwise.} \end{cases}$$