

Spectral bounds for mixing times

When working with finite-state, irreducible, aperiodic, reversible Markov chains, we can use the eigenvalues of the transition matrix to bound mixing times. These are some notes I took during an independent study with Professor Ursula Porod.

Spectral decomposition

First we note that the stationary distribution π is strictly positive because the Markov chain is irreducible. Therefore, we define

$$\mathbf{D} := \text{diag}(\pi(1), \dots, \pi(n))$$

and

$$\mathbf{P}^* := \mathbf{D}^{\frac{1}{2}} \mathbf{P} \mathbf{D}^{-\frac{1}{2}}.$$

Then

$$P_{ij}^* = \frac{\sqrt{\pi(i)}}{\sqrt{\pi(j)}} P_{ij}.$$

This follows directly from the computation of \mathbf{P}^* . Since \mathbf{P}^* is assumed to be reversible, we know $P_{ij}^* = P_{ji}^*$. Hence \mathbf{P}^* is symmetric, and its eigenvectors are orthonormal by the real spectral theorem. It shares the same eigenvalues with \mathbf{P} since they are similar matrices. Moreover, since our Markov chain is assumed to be irreducible and aperiodic, we know $\mathbf{P}^* s_1 = s_1$ where s_1 is the stationary distribution corresponding to \mathbf{P}^* , and the rest of the eigenvalues $\lambda_2, \dots, \lambda_k$ are all in the range $(-1, 1]$. Let \mathbf{S} be the matrix of eigenvectors of \mathbf{P}^* i.e. each column is an eigenvector. Defining $\mathbf{\Lambda}$ as the diagonal matrix of eigenvalues, we have

$$\mathbf{P}^* = \mathbf{S} \mathbf{\Lambda} \mathbf{S}^T$$

since the inverse of \mathbf{S} is its transpose (it is orthonormal). It follows that

$$\mathbf{P}_{ij}^* = \sum_{k=1}^n S_{ik} \lambda_k S_{jk},$$

and then

$$\mathbf{P}_{ij} = \frac{\sqrt{\pi(j)}}{\sqrt{\pi(i)}} \sum_{k=1}^n S_{ik} \lambda_k S_{jk}.$$

Now, since

$$\mathbf{P} = \mathbf{D}^{-\frac{1}{2}} \mathbf{S} \mathbf{\Lambda} \mathbf{S}^T \mathbf{D}^{\frac{1}{2}},$$

we have

$$\mathbf{P}^m = \mathbf{D}^{-\frac{1}{2}} \mathbf{S} \mathbf{\Lambda}^m \mathbf{S}^T \mathbf{D}^{\frac{1}{2}},$$

and equivalently

$$P_{ij}^m = \frac{\sqrt{\pi(j)}}{\sqrt{\pi(i)}} \sum_{k=1}^n S_{ik} \lambda_k^m S_{jk}.$$

If we take the first term out of the sum since $\lambda_1 = 1$, we have

$$P_{ij}^m = \frac{\sqrt{\pi(j)}}{\sqrt{\pi(i)}} S_{i1} S_{j1} + \sum_{k=2}^n S_{ik} \lambda_k^m S_{jk}.$$

Note that the since $P_{ij}^m \xrightarrow{m \rightarrow \infty} \pi(j)$ and $\lim_{m \rightarrow \infty} \lambda^m = 0$, it follows that

$$P_{ij}^m = \pi(j) + \sum_{k=2}^n S_{ik} \lambda_k^m S_{jk}.$$

This equation tells us that mixing times depend on their non-trivial eigenvalues. Specifically, there exists constants for each $i, j \leq n$ such that

$$|P_{ij}^m - \pi(j)| \leq C_{ij} \lambda_*^m$$

where $\lambda_* := \max\{\lambda_2, \dots, \lambda_n\}$. Taking $C = \max_{i,j \leq n} C_{ij}$, we can conclude that for any initial distribution μ_0 , and $\mu_m := \mu_0 \mathbf{P}^m$, we have

$$\|\mu_m - \pi(j)\|_{TV} \leq C \lambda_*^m.$$