Eigenvalues and Markov Chains

Will Perkins

April 15, 2013

The Metropolis Algorithm

Say we want to sample from a different distribution, not necessarily uniform. Can we change the transition rates in such a way that our desired distribution is stationary?

Amazingly, yes. Say we have a distribution π over $\mathcal{X}$ so that

$$\pi(x) = \frac{w(x)}{\sum_{y \in \mathcal{X}} w(y)}$$

I.e. we know the proportions but not the normalizing constant (and $\mathcal{X}$ is much too big to compute it).

The Metropolis Algorithm

Metropolis-Hastings Algorithm

1. Create a graph structure on $\mathcal{X}$ so the graph is connected and has maximum degree D.
2. Define the following transition probabilities:
   1. $p(x, y) = \frac{1}{2D} (\max\{w(y)/w(x), 1\})$ if x and y are neighbors.
   2. $p(x, y) = 0$ if x and y are not neighbors
   3. $p(x, x) = 1 - \sum_{y \sim x} p(x, y)$
3. Check that this Markov chain is irreducible, aperiodic, reversible and has stationary distribution π.

Example

Say we want to sample large independent sets from a graph G. I.e.

$$P(I) = \frac{\lambda^{|I|}}{Z}$$

where $Z = \sum_J \lambda^{|J|}$ where the sum is over all independent sets. Note that this distribution gives more weight to the largest independent sets.

Use the Metropolis Algorithm to find a Markov Chain with this distribution as the stationary distribution.

Linear Algebra

Recall some facts from linear algebra:

- If A is a real symmetric, $n \times n$ matrix, then A has real eigenvalues and there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A.
- The eigenvalues of A^n are the eigenvalues of A raised to the n
- Rayleigh Quotient form of eigenvalues

Perron-Frobenius Theorem

Theorem

Let $A > 0$ be a matrix with all positive entries. Then there exists an eigenvalue $\lambda_0 > 0$ with eigenvector x_0 all of whose entries are positive so that

1. *If $\lambda \neq \lambda_0$ is another eigenvalue of A then $|\lambda| < \lambda_0$.*
2. *λ_0 has algebraic and geometric multiplicity 1*

Perron-Frobenius Theorem

Proof: Define a set of real numbers
$\Lambda = \{\lambda : Ax \geq \lambda x \text{ for some } x \geq 0\}$.
Show that $\Lambda \in [0, M]$ for some M. Then let $\lambda_0 = \max \Lambda$.

From the definition of Λ, there exists an $x_0 \geq 0$ so that $Ax_0 \geq \lambda_0 x_0$. Suppose $Ax_0 \neq \lambda x_0$. Then let $y = Ax_0$ and

$$A(y - \lambda_0 x_0) = Ay - \lambda_0 y > 0$$

since $A > 0$. But this is a contradiction. So $Ax_0 = \lambda_0 x_0$.

Perron-Frobenius Theorem

Now pick an eigenvalue $\lambda \neq \lambda_0$ with eigenvector x. Then

$$A|x| \geq |Ax| = |\lambda x| = |\lambda||x|$$

and so $|\lambda| \leq \lambda_0$.

Finally, we show that there is no other eigenvalue $|\lambda| = |\lambda_0|$. Consider $A_\delta = A = \delta I$ for small enough δ so the matrix is still positive. A_δ has eigenvalues $\lambda_0 - \delta$ and $\lambda - \delta$, and $|\lambda_0 - \delta| \geq |\lambda - \delta|$. But if $\lambda \neq \lambda_0$ is on the same circle in the complex plane as λ_0, this is a contradiction. [picture]

Perron-Frobenius Theorem

Finally, we address the multiplicity. Say x and y are linearly independent eigenvectors with eigenvalue λ_0. Then find α so that $x + \alpha\ y$ has non-negative entries, but at least one 0 entry. But since $A > 0$ and $A(x + \alpha y) = \lambda(x + \alpha y)$ there is a contradiction.

Application to Markov Chains

Check: the conclusions of the Perrron-Frobenius theorem hold for the transition matrix of a finite, aperiodic, irreducible Markov chain.

Rate of Convergence

Theorem

Consider the transition matrix P of a symmetric, aperiodic, irreducible Markov Chain on n states. Let μ be the uniform (stationary) distribution. Let $\lambda_1 = 1$ be the largest eigenvalue and λ_2 the second-largest in absolute values. Then

$$||\pi_m^{(x)} - \mu||_{TV} \leq \sqrt{n}|\lambda_2|^m$$

Proof: Start with the Jordan Canonical form of the matrix P. (A generalization of diagonalizing - we'll assume P is diagonalizable), i.e.

$$D = UPU^{-1}$$

The rows of U are the left eigenvectors of P and the columns of U^{-1} are the right eigenvectors.

Rate of Convergence

Order the eigenvalues $1 = \lambda_1 > |\lambda_2| > \dots$. The left eigenvector of λ_1 is the stationary distribution vector. The first right eigenvector is the all 1's vector.

Now write $P^n = U^{-1} D^n U$.
Write π_0 is the eigenvector basis:

$$\pi_0 = \mu + c_2 u_2 + \dots c_n u_n$$

and

$$\pi_m = \pi_0 P^m = \mu + \sum_{j=2}^{n} c_j \lambda_j^m u_j$$

where $|\lambda_j| \leq |\lambda_2| < 1$.

Eigenvalues of Graphs

The adjacency matrix A of a graph G is the matrix whose i, jth entry is 1 if $(i, j) \in E(G)$. The normalized adjacency matrix turns this into a stochastic matrix - for example, if G is d-regular, we divide A by d.

For d-regular graph, with normalized adjancey matrix A,

- What is λ_1?
- What does A correspond to in terms of Markov Chains?
- What does it mean if $\lambda_2 = 1$?
- What does it mean if $\lambda_n = -1$?

Cheeger's Inequality

For a d-regular graph, define the edge expansion of a cut $S \subset V$ as:

$$h(S) = \frac{|E(S, S^c)|}{d \min\{|S|, |S^c|\}}$$

The edge expansion of a graph G is

$$h(G) = \min_{S \subset V} h(S)$$

Cheeger's Inequality

Theorem (Cheeger's Inequality)

Let $1 = \lambda_1 \geq \lambda_2 \geq \dots$ *be the eigenvalues of the random walk on the* d*-regular graph* G*. Then*

$$\frac{1 - \lambda_2}{2} \leq h(G) \leq \sqrt{2(1 - \lambda_2)}$$

What does this say about mixing times of random walks on graphs?

Ehrenfest Urn

What are the eigenvalues and eigenvectors of the Ehrenfest Urn?