Mathematica Pannonica
17/2 (2006), 183–193

ON THE FUNDAMENTAL MATRIX OF FINITE STATE MARKOV CHAINS, ITS EIGENSYSTEM AND ITS RELATION TO HITTING TIMES

Christiane **Takacs**

Department of Stochastics, University of Linz, A-4040 Linz, Austria

Received: June 2006

MSC 2000: 60 J 10; 62 M 15

Keywords: Markov chain, fundamental matrix, hitting time matrix, hitting time identities, spectral clustering.

Abstract: For a finite state reversible and ergodic Markov chain we prove an intimate relationship between its fundamental matrix and its hitting time matrix. From this we derive hitting time identities. Relating the eigensystem of the fundamental matrix to the eigensystem of the transition matrix yields a new characterization of equivalence classes of states indicated by piecewise constant eigenvectors. Since the latter are used for spectral clustering the paper gives a hitting time interpretation for the resulting clusters.

Let $(X_n)_{n\in\mathbb{N}_0}$ be a Markov Chain with finite state space $\{1,...,N\}$ and transition matrix $P=(p_{ij})$, where $p_{ij}=\Pr[X_{n+1}=j|X_n=i]$. Assume P to be aperiodic, irreducible and reversible, i.e., there is a normalized row vector π with

$$\pi_i p_{ij}=\pi_j p_{ji}.$$

Because of

$$(\pi P)_j=\sum_{i=1}^{N}\pi_i p_{ij}=\pi_j\sum_{i=1}^{N}p_{ji}=\pi_j,$$

π is the density of a stationary distribution of the Markov chain. By

E-mail address: christiane.takacs@jku.at

the convergence theorem for finite Markov chains π corresponds to the unique distribution of $\lim_{n\to\infty} X_n$, and for all i we have $\pi_i > 0$. Thus the matrix $D = \text{diag}(\pi)$ is welldefined and invertible.

We denote by $\mathbf{1}$ the columnvector with all entries equal to 1 and note that then $\mathbf{1}\mathbf{1}^t$ is a matrix with all entries equal to 1. Further we denote by $\Pi = \mathbf{1}\mathbf{1}^t D$ the matrix whose rows are all equal to π. Note that the convergence theorem cited above is equivalent to

$$\lim_{n\to\infty} P^n = \Pi.$$

1. Fundamental matrix and expected hitting times

Following [1] we introduce the fundamental matrix Z of P by

$$Z = (I - (P - \Pi))^{-1}.$$

From Th. 6.1 of Ch. 6 in [1] we learn that

$$Z = I + \sum_{n=1}^{\infty} (P - \Pi)^n = I + \sum_{n=1}^{\infty} (P^n - \Pi).$$

Thus the fundamental matrix in a certain sense describes how the process started in some initial state differs from the stationary process.

Let $\sigma_j = \min\{n \geq 0 : X_n = j\}$ be the first hitting time of j and denote by $s_{ij} = \mathrm{E}_i\sigma_j$ the expected hitting time of j from i, further let $\tau_j = \min\{n \geq 1 : X_n = j\}$ be the first return time to j and denote by $t_{ij} = \mathrm{E}_i\tau_j$ the expected return time to j from i. Denote by $S = (s_{ij})$ the matrix of the expected hitting times. Note that $s_{ii} = 0$, $t_{ii} = 1/\pi_i$ and $s_{ij} = t_{ij}$ for $i \neq j$. The expected hitting time of j resp. return time to j from stationary start is denoted by $\mathrm{E}_\pi\sigma_j$ resp. $\mathrm{E}_\pi\tau_j$.

In [1] a hitting time interpretation is given for the diagonal of Z. The following theorem extends this interpretation to all of Z.

Theorem 1.1. *The fundamental matrix satisfies*

$$Z = \left(\mathbf{1}\mathbf{1}^t + \Pi S - S\right) D,$$

for its entries we have

$$z_{ij} = \pi_j \left(1 + \mathrm{E}_\pi\sigma_j - \mathrm{E}_i\sigma_j\right) = \pi_j \left(\mathrm{E}_\pi\tau_j - \mathrm{E}_i\sigma_j\right).$$

Note that for the diagonal elements of Z we have (compare Ex. 6.1 of [1])

$$z_{jj} = \pi_j \mathrm{E}_\pi\tau_j.$$

Proof. From first-step analysis we have

$$t_{ij} = 1 + \sum_{k \neq j} p_{ik} t_{kj} = 1 + \sum_{k} p_{ik} s_{kj}$$

and thus

$$s_{ij} = t_{ij} - \delta_{ij} t_{ii} = 1 + \sum_{k} p_{ik} s_{kj} - \frac{\delta_{ij}}{\pi_i}.$$

Thus we successively calculate

$$S = \mathbf{1}\mathbf{1}^t + PS - D^{-1},$$

$$(I - P)\, S = \mathbf{1}\mathbf{1}^t - D^{-1},$$

$$Z\,(I - P)\, S = Z\left(\mathbf{1}\mathbf{1}^t - D^{-1}\right).$$

Using $\Pi P = \Pi$ and the definition of Z we simplify the left side

$$\begin{aligned} Z\,(I - P) = Z - \left(I + \sum_{n=1}^{\infty} (P^n - \Pi)\right) P = \\ = I + \sum_{n=1}^{\infty} (P^n - \Pi) - \left(P + \sum_{n=2}^{\infty} (P^n - \Pi)\right) = \\ = I - \Pi. \end{aligned}$$

Together with $Z\mathbf{1} = \mathbf{1}$ this reduces the above matrix equation to

$$(I - \Pi)\, S = \mathbf{1}\mathbf{1}^t - ZD^{-1}$$

and we arrive at

$$Z = \left(\mathbf{1}\mathbf{1}^t + \Pi S - S\right) D$$

proving our first claim.

For the entries of Z we have

$$z_{ij} = \pi_j \left(1 + \sum_{k=1}^{N} \pi_k s_{kj} - s_{ij}\right) = \pi_j \left(1 + \mathrm{E}_\pi \sigma_j - \mathrm{E}_i \sigma_j\right).$$

The second claim is a consequence of

$$\mathrm{E}_\pi \tau_j = \sum_{k=1}^{N} \pi_k t_{kj} = \sum_{k \neq j} \pi_k t_{kj} + \pi_j t_{jj} = \sum_{k=1}^{N} \pi_k s_{kj} + 1. \ \Diamond$$

The above theorem tells us how to calculate Z from S. Turning the equation around gives S from Z (compare Thms. 6.3 and 6.4 of [1]).

Theorem 1.2. *The expected hitting time matrix S has the representation*

$$S = \left(\mathbf{1}\mathbf{1}^t \mathrm{diag}\ (Z) - Z\right) D^{-1},$$

for its entries we have

$$s_{ij} = \frac{z_{jj} - z_{ij}}{\pi_j}.$$

Proof. From the preceding theorem we have

$$z_{ij} = \pi_j \left(1 + \mathrm{E}_\pi \sigma_j - s_{ij}\right)$$
$$z_{jj} = \pi_j \left(1 + \mathrm{E}_\pi \sigma_j\right).$$

By substraction we arrive at

$$s_{ij} = \frac{z_{jj} - z_{ij}}{\pi_j}. \ \Diamond$$

Both theorems establish the very close relationship between the fundamental matrix and the matrix of hitting times.

2. The fundamental matrix is reversible and admits hitting time identities

First we prove detailled balance equations for Z. These, as a corollary, produce identities for certain hitting times.

Theorem 2.1. *For aperiodic, irreducible and reversible Markov chains with finite state space we have*

$$\pi_i z_{ij} = \pi_j z_{ji}.$$

Proof. Since $Z = I + \sum_{n=1}^{\infty} (P^n - \Pi)$, i.e.,

$$z_{ij} = \delta_{ij} + \sum_{n=1}^{\infty} \left(p_{ij}^{(n)} - \pi_j\right)$$

the assertion is a consequence of $\pi_i \delta_{ij} = \pi_j \delta_{ji}$ and

$$\pi_i \left(p_{ij}^{(n)} - \pi_j\right) = \pi_j \left(p_{ji}^{(n)} - \pi_i\right) \text{ for } n \in \mathbb{N}$$

which is proved by a standard induction argument. $\Diamond$

Corollary 2.2. *For aperiodic, irreducible and reversible Markov chains with finite state space we have*

$$\mathrm{E}_\pi \sigma_j + \mathrm{E}_j \sigma_i = \mathrm{E}_\pi \sigma_i + \mathrm{E}_i \sigma_j,$$
$$\mathrm{E}_\pi \tau_j + \mathrm{E}_j \tau_i = \mathrm{E}_\pi \tau_i + \mathrm{E}_i \tau_j.$$

Proof. Prompting $z_{ij} = \pi_j \left(1 + \mathrm{E}_\pi \sigma_j - \mathrm{E}_i \sigma_j\right)$ into

$$\pi_i z_{ij} = \pi_j z_{ji}$$

we have

$$\mathrm{E}_\pi \sigma_j - \mathrm{E}_i \sigma_j = \mathrm{E}_\pi \sigma_i - \mathrm{E}_j \sigma_i$$

proving our first claim.

Because of $\mathrm{E}_\pi \sigma_j = \mathrm{E}_\pi \tau_j - 1$ and $\mathrm{E}_i \sigma_j = \mathrm{E}_i \tau_j - \delta_{ij} \mathrm{E}_i \tau_i$ our second claim is equivalent to the first. ◊

According to the corollary the stationary Markov chain expectedly needs the same time to visit state i after j, as it would take to pass i (regardless of having hit j) and to visit j afterwards. This feature is not obvious from time reversal.

3. The eigensystem of the fundamental matrix

The following section continues to explore the relationship between Z and P.

Since $D = \mathrm{diag}\ (\pi)$ is invertible, the matrix $D^{1/2} P D^{-1/2}$ is welldefined and because of reversibility it is a symmetric matrix. Thus it is diagonalizable and P is diagonalizable, too. If $D^{1/2} v$ is a right eigenvector of $D^{1/2} P D^{-1/2}$ corresponding to the eigenvalue λ, the vector v is a right eigenvector of P corresponding to the same eigenvalue. The corresponding left eigenvector of $D^{1/2} P D^{-1/2}$ is $v^t D^{1/2}$, the left eigenvector of P corresponding to v is $v^t D$.

Thus we have the representation

$$P = V \Lambda V^t D,$$

where V is the collection of right eigenvectors of P such that $V^t D V = I$, and Λ is the diagonal matrix of the eigenvalues corresponding to the columns of V.

Proposition 3.1. *Let P be an aperiodic, irreducible and reversible transition matrix with stationary distribution π and $D = \mathrm{diag}\ (\pi)$, let V be the collection of right eigenvectors of P, such that $V^t D V = I$ and Λ the diagonal matrix of the eigenvalues corresponding to the columns of V. Then we have the representation*

$$Z = V \tilde{\Lambda} V^t D,$$

where $\tilde{\lambda}_i = (1 - \lambda_i)^{-1}$ for $\lambda_i \neq 1$ and $\tilde{\lambda}_i = 1$ for $\lambda_i = 1$.

Proof. Without loss of generality let $V = (v_1, \ldots, v_N)$ with $v_1 = \mathbf{1}$. Then we have $V^t D = (v_1^t D, \ldots, v_N^t D)^t$ with $v_1^t D = \pi$, and $\Lambda = \mathrm{diag}\ (\lambda_1, \ldots, \lambda_N)$ with $\lambda_1 = 1$ and $\lambda_i \neq 1$ for $i \neq 1$.

With $Z^{-1} = I - (P - \Pi)$ we have

$$V^t D Z^{-1} V = V^t D V - V^t D (P - \Pi) V = I - V^t D P V + V^t D \Pi V,$$

where $V^t D P V = \Lambda = \mathrm{diag}\ (1, \ldots, \lambda_N)$ and

$$V^tD\Pi V = \left(\pi, \dots, v_N^tD\right)^t (\pi, \dots, \pi)^t (\mathbf{1}, \dots, v_N) =$$
$$= \left(\pi, \dots, v_N^tD\right)^t (\mathbf{1}, o, \dots, o) = \operatorname{diag}(1, 0, \dots, 0).$$

We arrive at

$$V^tDZ^{-1}V = I - \operatorname{diag}(1, \dots, \lambda_N) + \operatorname{diag}(1, 0, \dots, 0) =$$
$$= \operatorname{diag}(1, 1 - \lambda_2, \dots, 1 - \lambda_N).$$

Solving for Z leads to the representation

$$Z = \left(\left(V^tD\right)^{-1} \operatorname{diag}(1, 1 - \lambda_2, \dots, 1 - \lambda_N) V^{-1}\right)^{-1} =$$
$$= V \operatorname{diag}\left(1, (1 - \lambda_2)^{-1}, \dots, (1 - \lambda_N)^{-1}\right) V^tD,$$

which proves our claim. ◊

Note that according to the above proposition the fundamental matrix Z is diagonalizable, where the eigenvectors of Z coincide with those of P and for each eigenvalue $\tilde{\lambda}$ of Z corresponding to an eigenvector with eigenvalue λ of P we have

$$\tilde{\lambda} = \begin{cases} \frac{1}{1-\lambda} & \text{for } \lambda \neq 1 \\ 1 & \text{for } \lambda = 1 \end{cases}.$$

4. Characterizations of equivalence classes

An interesting question for Markov chains is, whether states may be agglomerated to classes of states such that the stochastic process induced on the classes is again a Markov chain. It is well known that agglomeration is possible if for all states in an arbitrary class the transition probabilities to the other classes are the same. To be precise, if $(C_1, \dots, C_K)$ is a partition of the states and if for all $r, s = 1, \dots, K$ we have

$$\sum_{\sigma \in C_s} p_{i\sigma} = \sum_{\sigma \in C_s} p_{j\sigma} \text{ for all } i, j \in C_r, \tag{4.1}$$

then the stochastic process with state space $(C_1, \dots, C_K)$, which arises from the original Markov chain by observing the classes of the states, is again a Markov chain and has transition probabilities

$$\bar{p}_{rs} = P\left(X_{n+1} \in C_s \mid X_n \in C_r\right) = \sum_{\sigma \in C_s} p_{i\sigma} \text{ with } i \in C_r.$$

In this situation we say that the Markoc chain (resp. P) admits agglomeration with respect to the partition $(C_1, \ldots, C_K)$ and call the Markov chain with transition matrix $\bar{P} = (\bar{p}_{rs})$ the agglomerated Markov chain.

Proposition 4.1. *For an aperiodic, irreducible and reversible P the agglomerated Markov chain is aperiodic, irreducible and reversible, too.*

Proof. Since the period of a state σ is the greatest common divisor of the set $\{n : p_{\sigma\sigma}^{(n)} > 0\}$, which for $\sigma \in C_s$ is a subset of $\{n : \bar{p}_{ss}^{(n)} > 0\}$, the period of $\bar{P}$ cannot be longer than the period of P. So $\bar{P}$ is aperiodic.

Since the original chain is irreducible and $p_{\rho\sigma}^{(n)} \le \bar{p}_{rs}^{(n)}$ where $\rho \in C_r$ and $\sigma \in C_s$, the agglomerated Markov chain is irreducible, too.

Since P is reversible we have

$$\sum_{\rho \in C_r} \sum_{\sigma \in C_s} \pi_\rho p_{\rho\sigma} = \sum_{\sigma \in C_s} \sum_{\rho \in C_r} \pi_\sigma p_{\sigma\rho}.$$

Now the left side evaluates to

$$\sum_{\rho \in C_r} \sum_{\sigma \in C_s} \pi_\rho p_{\rho\sigma} = \sum_{\rho \in C_r} \pi_\rho \sum_{\sigma \in C_s} p_{\rho\sigma} = \Big(\sum_{\rho \in C_r} \pi_\rho \Big) \bar{p}_{rs}$$

and the right side to $\left(\sum_{\sigma \in C_s} \pi_\sigma\right) \bar{p}_{sr}$. With $\bar{\pi}_s = \sum_{\sigma \in C_s} \pi_\sigma$ we have proved

$$\bar{\pi}_r \bar{p}_{rs} = \bar{\pi}_s \bar{p}_{sr},$$

and thus reversibility of the agglomerated Markov chain. ◊

Note that $\bar{\pi}_s = \sum_{\sigma \in C_s} \pi_\sigma$ is the stationary distribution of the agglomerated chain. As a consequence of the above proposition $\bar{P}$ is diagonalizable and has a representation analogous to P, i.e., $\bar{P} = \bar{V} \bar{\Lambda} \bar{V}^t \bar{D}$.

For reversible chains the agglomeration condition (4.1) can be rewritten as

$$\frac{1}{\pi_i} \sum_{\sigma \in C_s} \pi_\sigma p_{\sigma i} = \frac{1}{\pi_j} \sum_{\sigma \in C_s} \pi_\sigma p_{\sigma j} \text{ for all } i, j \in C_r,$$

i.e., for the stationary chain starting in one class, the probability of a step to a state in some (other) class is proportional to the stationary probability of the state, where the factor of proportionality only depends on the classes. Thus the stationary transition probability from a class to a state in some (other) class has a product form, where the first factor describes the transition probability between the classes and the second factor the stationary probability of the state within the class,

$$P\left(X_{n+1} = i \mid X_n \in C_s\right) = \frac{1}{\bar{\pi}_s} \sum_{\sigma \in C_s} \pi_\sigma p_{\sigma i} = \bar{p}_{sr} \frac{\pi_i}{\bar{\pi}_r} \text{ for } i \in C_r.$$

In our setting the agglomeration condition (4.1) can also be rewritten in terms of the eigenvectors of P. To do so, we need Th. 4.4 of [5], which dates back to [3]. We cite the theorem here for convenience.

Proposition 4.2. *Let $M \in \mathbb{R}^n_n$ be a matrix and $(C_1, \dots, C_K)$ a partition of its index set $\{1, \dots, N\}$. Then we have:*

M has the eigenvalues $\lambda_1, \dots, \lambda_K$ and the corresponding right eigenvectors are linearly independent and piecewise constant on $(C_1, \dots, C_K)$ iff for all $r, s = 1, \dots, K$

$$\sum_{\sigma \in C_s} m_{i\sigma} = \sum_{\sigma \in C_s} m_{j\sigma} \text{ for all } i, j \in C_r$$

and

$$(\bar{m}_{rs})_{r,s=1,\dots,K} \text{ where } \bar{m}_{rs} = \sum_{\sigma \in C_s} m_{i\sigma} \text{ for } i \in C_r$$

can be diagonalized with the same eigenvalues $\lambda_1, \dots, \lambda_K$.

Theorem 4.3. *An aperiodic, irreducible and reversible transition matrix P admits agglomeration with respect to the partition $(C_1, \dots, C_K)$ iff it has K linearly independent right eigenvectors, which are piecewise constant on $(C_1, \dots, C_K)$. The corresponding eigenvalues coincide with the eigenvalues of the transition matrix of the agglomerated chain.*

Proof. Since the agglomeration condition (4.1) implies $\bar{P}$ to be diagonalizable the assertion is immediate from the above proposition. ◊

Since the eigenvectors of Z coincide with the eigenvectors of P, we can use the above proposition to derive an agglomeration condition in terms of Z.

Theorem 4.4. *Given an aperiodic, irreducible and reversible transition matrix P and a partition $(C_1, \dots, C_K)$, then P admits agglomeration with respect to the partition iff one of the following is true:*

- *Z has K linearly independent right eigenvectors, which are piecewise constant on $(C_1, \dots, C_K)$.*
- *For the fundamental matrix Z we have for all $r, s = 1, \dots, K$*

$$\sum_{\sigma \in C_s} z_{i\sigma} = \sum_{\sigma \in C_s} z_{j\sigma} \text{ for all } i, j \in C_r. \tag{4.2}$$

In any case

$$\bar{Z} = (\bar{z}_{rs})_{r,s=1,\dots,K} \text{ where } \bar{z}_{rs} = \sum_{\sigma \in C_s} z_{i\sigma} \text{ for } i \in C_r$$

can be diagonalized, where the eigenvectors coincide with the eigenvectors of the transition matrix $\bar{P}$ of the agglomerated chain, and the eigenvalues are obtained by

$$\tilde{\lambda} = \begin{cases} \frac{1}{1-\lambda} & \text{for } \lambda \neq 1 \\ 1 & \text{for } \lambda = 1 \end{cases},$$

where λ is an eigenvalue of $\bar{P}$.

Proof. The equivalence of the first assertion is a consequence of Z and P having the same eigenvectors. By Prop. 4.2 the first assertion implies the second. Now suppose the second assertion is true. From $\pi_i z_{ij} = = \pi_j z_{ji}$ and condition (4.2) we conclude $\bar{\pi}_r \bar{z}_{rs} = \bar{\pi}_s \bar{z}_{sr}$ analogously to Prop. 4.1. Thus $\bar{D}^{1/2}\bar{Z}\bar{D}^{-1/2}$ is a symmetric matrix and consequently $\bar{Z}$ is diagonalizable. This together with condition (4.2) is enough to conclude the first assertion by Prop. 4.2. Thus both assertions are equivalent. Our claim concerning the eigenvectors and eigenvalues follows from the representations of Z and $\bar{Z}$. $\diamond$

It is interesting to interpret the agglomeration condition for Z (4.2) in terms of hitting times. By $\mathrm{E}_{\pi(\cdot|C_s)}$ we denote the expectation with respect to a stationary initial distribution conditioned to start in class C_s.

Corollary 4.5. *The agglomeration condition for Z (4.2) is equivalent to each of the following conditions, where $r, s = 1, \ldots, K$ and $i, j \in C_r$*

$$\sum_{\sigma \in C_s} \pi_\sigma \mathrm{E}_i \sigma_\sigma = \sum_{\sigma \in C_s} \pi_\sigma \mathrm{E}_j \sigma_\sigma,$$

$$\mathrm{E}_{\pi(\cdot|C_s)}(\sigma_i - \sigma_j) = \mathrm{E}_\pi(\sigma_i - \sigma_j).$$

Proof. Using the definition of Z gives the first condition via

$$\sum_{\sigma \in C_s} z_{i\sigma} = \sum_{\sigma \in C_s} z_{j\sigma},$$

$$\sum_{\sigma \in C_s} \pi_\sigma (1 + \mathrm{E}_\pi \sigma_\sigma - \mathrm{E}_i \sigma_\sigma) = \sum_{\sigma \in C_s} \pi_\sigma (1 + \mathrm{E}_\pi \sigma_\sigma - \mathrm{E}_j \sigma_\sigma),$$

$$\sum_{\sigma \in C_s} \pi_\sigma \mathrm{E}_i \sigma_\sigma = \sum_{\sigma \in C_s} \pi_\sigma \mathrm{E}_j \sigma_\sigma.$$

By reversibility of Z this is equivalent to

$$\frac{1}{\pi_i} \sum_{\sigma \in C_s} \pi_\sigma z_{\sigma i} = \frac{1}{\pi_j} \sum_{\sigma \in C_s} \pi_\sigma z_{\sigma j},$$

$$\frac{1}{\pi_i} \sum_{\sigma \in C_s} \pi_\sigma \pi_i (1 + \mathrm{E}_\pi \sigma_i - \mathrm{E}_\sigma \sigma_i) = \frac{1}{\pi_j} \sum_{\sigma \in C_s} \pi_\sigma \pi_j (1 + \mathrm{E}_\pi \sigma_j - \mathrm{E}_\sigma \sigma_j),$$

$$\sum_{\sigma \in C_s} \pi_\sigma (\mathrm{E}_\pi \sigma_i - \mathrm{E}_\sigma \sigma_i) = \sum_{\sigma \in C_s} \pi_\sigma (\mathrm{E}_\pi \sigma_j - \mathrm{E}_\sigma \sigma_j),$$

$$\sum_{\sigma\in C_s} \pi_\sigma \left(\mathrm{E}_\pi \sigma_i - \mathrm{E}_\pi \sigma_j\right) = \sum_{\sigma\in C_s} \pi_\sigma \left(\mathrm{E}_\sigma \sigma_i - \mathrm{E}_\sigma \sigma_j\right),$$

$$\mathrm{E}_\pi \left(\sigma_i - \sigma_j\right) = \frac{1}{\pi\left(C_s\right)} \sum_{\sigma\in C_s} \pi_\sigma \left(\mathrm{E}_\sigma \sigma_i - \mathrm{E}_\sigma \sigma_j\right),$$

$$\mathrm{E}_\pi \left(\sigma_i - \sigma_j\right) = \mathrm{E}_{\pi(\cdot|C_s)} \left(\sigma_i - \sigma_j\right) \ \Diamond$$

Note that the first characterization may be written as

$$\sum_{\sigma\in C_s} \frac{\mathrm{E}_i \tau_\sigma}{\mathrm{E}_\sigma \tau_\sigma} = \sum_{\sigma\in C_s} \frac{\mathrm{E}_j \tau_\sigma}{\mathrm{E}_\sigma \tau_\sigma} \text{ for } i, j \in C_r.$$

The second characterization means that the expected difference of the hitting times of states in one class is the same for the stationary Markov chain and for the stationary Markov chain conditioned to start in some class.

5. Application to clustering

Since the eigenvectors of Z are the same as those of P, there is an interesting interpretation in the context of spectral segmentation, where one wants to identify clusters of points, i.e., group together points which are near or similar. In the literature there are several algorithms (see [2], [4], [6] for a survey) finding an optimal partition. Spectral methods have in common that they use piecewise almost constant eigenvectors as indicators for the underlying partition. Frequently the eigenvectors of a transition matrix P (which arises from some similarity matrix by normalization of the rows) are used. By perturbation arguments the clustering properties can be interpreted in terms of agglomeration of states of a Markov chain, where the agglomeration conditions (4.1) and (4.2) explain which kind of similarity is used for the classification of states. While a classification with respect to P is based on one step transition probabilities, a classification with respect to Z is based on hitting time similarities. The key observation is that both classifications coincide. Maybe this is another reason why spectral methods based on P (resp. Z) are especially powerful.

References

[1] BRÉMAUD, P.: Markov Chains Gibbs Fields, Monte Carlo Simulation and Queues, Springer, New York, 1999.

[2] KANNAN, R., VEMPALA, S. and VETTA, A.: On Clusterings – Good, Bad and Spectral, Proceedings of 41^{st} Annual Symposium on Foundations of Computer Science, 2000, 367–377.

[3] MEILA, M. and SHI, J.: A Random Walks View of Spectral Segmentation, Proceedings of the International Workshop on AI and Statistics, 2001.

[4] SEARY, A. and RICHARDS, W.: Spectral Methods for Analyzing and Visualizing Networks: An Introduction, Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers, 2003.

[5] TAKACS, C.: Partitioning with Piecewise Constant Eigenvectors, Mathematics and Computer Science III: Algorithms, Trees, Combinatorics and Probabilities, Birkhäuser, 2004, 363–374.

[6] WEISS, Y.: Segmentation using eigenvectors: A unifying view, International Conference on Computer Vision, 1999.