Optimization and Computational Linear Algebra for Data Science
Lecture 6: Eigenvalues, eigenvectors and Markov chains

Léo Miolane · leo.miolane@gmail.com

September 30, 2020

Warning: *This material is not meant to be lecture notes. It only gathers the main concepts and results from the lecture, without any additional explanation, motivation, examples, figures...*

1 Eigenvalues and eigenvectors

Definition 1.1

Let $A \in \mathbb{R}^{n \times n}$. A **non-zero** vector $v \in \mathbb{R}^n$ is said to be an eigenvector of A is there exists $\lambda \in \mathbb{R}$ such that
$$Av = \lambda v.$$
The scalar λ is called the eigenvalue (of A) associated to v. The set
$$E_\lambda(A) = \{x \in \mathbb{R}^n \,|\, Ax = \lambda x\} = \mathrm{Ker}(A - \lambda \mathrm{Id})$$
is called the eigenspace of A associated to λ. The dimension of $E_\lambda(A)$ is called the multiplicity of the eigenvalue λ.

Remark 1.1. *Notice that $E_\lambda(A)$ is a subspace of $\mathbb{R}^n$: any (non-zero) linear combination of eigenvectors associated with the eigenvalue λ is also an eigenvector of A associated with λ.*

Remark 1.2. *Definition 1.1 can be generalized to allow complex eigenvalues and eigenvectors: $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n$. However in this course, we only consider **real eigenvalues and eigenvectors**.*

Example 1.1. For $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, we introduce the notation

$$\mathrm{Diag}(\lambda_1, \ldots, \lambda_n) \overset{\mathrm{def}}{=} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Let $(e_1, e_2, \ldots, e_n)$ be the canonical basis of $\mathbb{R}^n$. Then, for all $i \in \{1, \ldots, n\}$

$$\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)\, e_i = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \lambda_i & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_i e_i.$$

Hence the vector e_i is an eigenvector of the matrix $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ associated with the eigenvalue λ_i.

Proposition 1.1

Let $A \in \mathbb{R}^{n \times n}$. Suppose that A has an eigenvalue $\lambda \in \mathbb{R}$ and let $x \in \mathbb{R}^n$ be an eigenvector associated to λ. The following holds:

- For all $\alpha \in \mathbb{R}$, $\alpha\lambda$ is an eigenvalue of the matrix αA and x is an associated eigenvector.

- For all $\alpha \in \mathbb{R}$, $\lambda + \alpha$ is an eigenvalue of the matrix $A + \alpha\mathrm{Id}$ and x is an associated eigenvector.

- For all $k \in \mathbb{N}$, λ^k is an eigenvalue of the matrix A^k and x is an associated eigenvector.

- If A is invertible then $1/\lambda$ is an eigenvalue of the matrix inverse A^{-1} and x is an associated eigenvector.

Definition 1.2

The set of all eigenvalues of A is called the spectrum of A and denoted by $\mathrm{Sp}(A)$.

Proposition 1.2

Let $v_1, \ldots, v_k$ be eigenvectors of A corresponding (respectively) to the eigenvalues $\lambda_1, \ldots, \lambda_k$. If the λ_i are all distinct ($\lambda_i \neq \lambda_j$ for all $i \neq j$) then the vectors $v_1, \ldots, v_k$ are linearly independent.

It follows from Proposition 1.2:

Corollary 1.1

A $n \times n$ matrix A admits at most n different eigenvalues: $\#\mathrm{Sp}(A) \leq n$.

In fact, we have the following stronger result:

Proposition 1.3

Let $A \in \mathbb{R}^{n \times n}$. If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of A of multiplicities $m_1, \ldots, m_k$ respectively, then

$$m_1 + \cdots + m_k \leq n.$$

2 Application to Markov chains

2.1 First definitions and properties

A finite Markov chain is a random process which moves among the elements of a finite set E in the following manner: when at $x \in E$, the next position is chosen according to a fixed probability distribution $P(x, \cdot)$. More formally:

Definition 2.1

A sequence of random variables $(X_0, X_1, \ldots)$ is a Markov chain with state space E and "transition matrix" P if for all $t \geq 0$,

$$\mathbb{P}(X_{t+1} = y \mid X_0 = x_0, \ldots, X_t = x_t) = P(x_t, y)$$

for all $x_0, \ldots, x_t$ such that $\mathbb{P}(X_0 = x_0, \ldots, X_t = x_t) > 0$.

The transition matrix P verifies therefore, for all $x \in E$,

$$\sum_{y \in E} P(x, y) = 1. \tag{1}$$

In order to simplify the notations, we will assume that $E = \{1, 2, \ldots, n\}$ and write for all $i, j \in E$, $P_{i,j} = P(j, i)$. **Note that we switched here the order of i and j. This is not what is usually done in the literature, but this will allow us to be more coherent with our linear algebra framework**. Such matrix is said to be stochastic:

Definition 2.2 (*Stochastic matrix*)

A matrix $P \in \mathbb{R}^{n \times n}$ is said to be stochastic if:

(i) $P_{i,j} \geq 0$ for all $1 \leq i, j \leq n$.

(ii) $\sum\limits_{i=1}^{n} P_{i,j} = 1$, for all $1 \leq j \leq n$.

Let $(X_0, X_1, \ldots)$ be a Markov chain on $\{1, \ldots, n\}$ with transition matrix P. For $t \geq 0$ we will encode the distribution of X_t in the vector

$$x^{(t)} = (x_1^{(t)}, \ldots x_n^{(t)}) = \big(\mathbb{P}(X_t = 1), \ldots, \mathbb{P}(X_t = n)\big) \in \Delta_n$$

where Δ_n is the "n-simplex"

$$\Delta_n \overset{\text{def}}{=} \Big\{ x \in \mathbb{R}^n \ \Big| \ \sum_{i=1}^{n} x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i \Big\}.$$

Proposition 2.1

For all $t \geq 0$

$$x^{(t+1)} = P x^{(t)} \quad \text{and consequently,} \quad x^{(t)} = P^t x^{(0)}.$$

Proof. Let $i \in \{1, \ldots, n\}$.

$$x_i^{(t+1)} = \mathbb{P}(X_{t+1} = i) = \sum_{j=1}^{n} \mathbb{P}(X_{t+1} = i | X_t = j) \mathbb{P}(X_t = j) = \sum_{i=1}^{n} P_{i,j} x_j^{(t)} = (P x^{(t)})_i.$$

$\square$

Corollary 2.1

Let P be a stochastic matrix. Then

- *For all $x \in \Delta_n$, $Px \in \Delta_n$.*

- *For all $t \geq 1$, P^t is stochastic.*

2.2 Invariant measures and the Perron-Frobenius Theorem

We will be interested in the distribution of X_t for large t, that is the limit of $x^{(t)} = P^t x^{(0)}$. As we will see, under suitable conditions on the matrix P, $x^{(t)}$ converge to some $\mu \in \Delta_n$ as $t \to \infty$. In such case, by taking the $t \to \infty$ limit in $x^{(t+1)} = P x^{(t)}$ we get $\mu = P\mu$. This motivates the following definition:

Definition 2.3 (*Invariant measure*)

A vector $\mu \in \Delta_n$ is called an invariant measure for the transition matrix P if $\mu = P\mu$, i.e. if μ is an eigenvector of P associated with the eigenvalue 1.

Theorem 2.1 (*Perron-Frobenius, stochastic case*)

> Let P be a stochastic matrix such that there exists $k \geq 1$ such that all the entries of P^k are strictly positive. Then the following holds:
>
> (i) 1 is an eigenvalue of P and there exists an eigenvector $\mu \in \Delta_n$ associated to 1.
> (ii) The eigenvectors associated to 1 are unique up to scalar multiple (i.e. $\mathrm{Ker}(P - \mathrm{Id}) = \mathrm{Span}(\mu)$).
> (iii) For all $x \in \Delta_n$, $P^t x \xrightarrow[t \to \infty]{} \mu$.

Theorem 2.1 is proved in the next section. Theorem 2.1 tells us that there is a unique $\mu \in \Delta_n$ such that $P\mu = \mu$. We call μ the Perron-Frobenius eigenvector of P.

Remark 2.1. *There exist a stronger version of the Perron-Frobenius Theorem which does not require the columns of P to sum to 1, see for instance Theorem 1.1 in [2]. The proof is however more involved.*

Corollary 2.2

> Let P be a stochastic matrix such that there exists $k \geq 1$ such that all the entries of P^k are strictly positive. Then there exists a unique invariant measure μ and for all initial condition $x^{(0)} \in \Delta_n$,
> $$x^{(t)} = P^t x^{(0)} \xrightarrow[t \to \infty]{} \mu.$$

Corollary 2.2 tells us that the Markov chain "forgets" its initial condition to converge to its invariant measure μ. We say that the chain is "mixing".

Working a little bit more, one can prove the "ergodic" Theorem that states that μ_i corresponds to the average time spent by the Markov chain in state i.

Theorem 2.2 (*Ergodic Theorem*)

> Let $(X_t)_{t \geq 0}$ be a Markov chain whose transition matrix is P. Assume that there exists $k \geq 1$ such that all the entries of P^k are strictly positive and let μ be the unique invariant measure of P. Then for any initial condition X_0, we have with probability 1, for all $i = 1, \ldots, n$:
> $$\frac{1}{T} \#\{t < T \mid X_t = i\} \xrightarrow[T \to \infty]{} \mu_i.$$

2.3 Proof of Theorem 2.1

We first prove the theorem in the case $k = 1$, when $P_{i,j} > 0$ for all i, j.

Lemma 2.1

> The mapping
> $$\varphi : \begin{array}{ccc} \Delta_n & \to & \Delta_n \\ x & \mapsto & Px \end{array}$$
> is a contraction mapping for the ℓ_1-norm: there exists $c \in (0, 1)$ such that for all $x, y \in \Delta_n$:
> $$\|Px - Py\|_1 \leq c\|x - y\|_1.$$

Proof. First notice that φ is well-defined by Corollary 2.1. Let us write $\alpha \stackrel{\text{def}}{=} \min_{i,j} P_{i,j} \in (0,1)$. Let $x, y \in \Delta_n$. We will show that $\|Px - Py\|_1 \leq (1-\alpha)\|x-y\|_1$, i.e. $\|Pz\|_1 \leq \alpha\|z\|_1$ where $z = x - y$. Compute

$$\|Pz\|_1 = \sum_{i=1}^n |(Pz)_i| = \sum_{i=1}^n \Big| \sum_{j=1}^n P_{i,j} z_j \Big|.$$

Since $\sum_j z_j = 0$ we have $\sum_j (P_{i,j} - \alpha/n) z_j = \sum_j P_{i,j} z_j$. Hence

$$\|Pz\|_1 = \sum_{i=1}^n \Big| \sum_{j=1}^n (P_{i,j} - \alpha/n) z_j \Big| \leq \sum_{i=1}^n \sum_{j=1}^n (P_{i,j} - \alpha/n)|z_j| = \sum_{j=1}^n (1-\alpha)|z_j| = (1-\alpha)\|z\|_1.$$

$\square$

Let now μ be a minimizer of $x \mapsto \|Px - x\|_1$ on Δ_n.

- $P\mu = \mu$, because otherwise $\|P\mu - \mu\|_1 > 0$ and by Lemma 2.1

$$\|P(P\mu) - P\mu\|_1 \leq c\|P\mu - \mu\|_1 < \|P\mu - \mu\|_1,$$

  which contradicts the optimality of μ. This proves (i).

- For all $x \in \Delta_n$, we have

$$\|P^t x - \mu\|_1 = \|P^t x - P^t \mu\|_1 \leq c^t \|x - \mu\|_1 \xrightarrow[t \to \infty]{} 0,$$

  which proves (iii).

- Let us now prove (ii). Let $x \in \mathbb{R}^n$ such that $Px = x$. Then for all $t \geq 1$:

$$x = P^t x = x_1 P^t e_1 + \cdots + x_n P^t e_n \tag{2}$$
$$\xrightarrow[t \to \infty]{} x_1 \mu + \cdots + x_n \mu = (x_1 + \cdots + x_n)\mu \in \text{Span}(\mu), \tag{3}$$

  which proves (ii).

This proves Theorem 2.1 in the case $k = 1$.

In the case $k > 1$ we simply apply the result for $k = 1$ to P^k. This gives that there exists a unique $\mu \in \Delta_n$ such that $P^k \mu = \mu$. Multiplying by P on both sides leads to $P^k(P\mu) = P\mu$. Since $P\mu \in \Delta_n$ we obtain that $P\mu = \mu$ by uniqueness of μ. This proves (i). To prove (ii) we consider $x \in \mathbb{R}^n$ such that $Px = x$. By iteration we get $P^k x = x$ which implies (using the result on P^k) that $x \in \text{Span}(\mu)$. To prove (iii) we fix $\ell \in \{0, \ldots, k-1\}$. Let $x \in \Delta_n$. By applying the point (iii) to P^k, we have

$$P^{kt} P^\ell x \xrightarrow[t \to \infty]{} \mu.$$

Since this holds for all $\ell \leq k - 1$ we obtain that $P^r x \xrightarrow[r \to \infty]{} \mu$ using the Euclidean division of r by k.

3 Example: Google's PageRank algorithm

3.1 The PageRank algorithm

The PageRank algorithm was invented by Larry Page and Sergey Brin [1]. The goal is to rank n web pages in terms of "importance". L. Page and S. Brin considered a "drunk surfer" that goes

from a page j to an other page i by randomly clicking on the links that are on the page j. This can be modeled by a Markov chain with state space $\{1, \ldots, n\}$ and transition matrix P given by

$$P_{i,j} = \begin{cases} 1/\deg(j) & \text{if there is a link } j \to i \\ 0 & \text{otherwise,} \end{cases}$$

where $\deg(j)$ denotes the number of outgoing links on page j. The idea behind PageRank is to quantify the importance of a page i by the fraction of time spent by the "drunk surfer" on it. By Theorem 2.2 we know that this corresponds to the coefficient μ_i of the invariant measure μ of P.

The matrix P is however not guaranteed to satisfy the hypotheses of Theorem 2.2-Corollary 2.2. Brin and Page proposed to use instead of P the matrix

$$G = \alpha P + \frac{1-\alpha}{n} \mathbb{1},$$

where $\alpha \in (0,1)$ is a parameter close to 1 (Google takes $\alpha \simeq 0.85$) and $\mathbb{1}$ denotes the all-one matrix. The PageRank algorithm computes μ the Perron-Frobenius eigenvector of the matrix G and ranked the webpages according to their coordinate in the vector μ: the higher μ_i, the better page i will be ranked.

3.2 Ranking tennis players

The ideas behind PageRank can be applied in many different contexts, not only for ranking webpages. In the following example we aim at ranking the following $n = 54$ tennis players:

Federer, Nadal, Djokovic, Murray, Del Potro, Roddick, Coria, Zverev, Ferrer, Soderling, Tsonga, Nishikori, Raonic, Nalbandian, Wawrinka, Berdych, Hewitt, Tsitsipas, Monfils, Gonzalez, Thiem, Ljubicic, Davydenko, Cilic, Pouille, Safin, Isner, Dimitrov, Medvedev, Ferrero, Goffin, Bautista Agut, Sock, Gasquet, Simon, Blake, Monaco, Coric, Stepanek, Khachanov, Almagro, Robredo, Verdasco, Anderson, Youzhny, Baghdatis, Dolgopolov, Kohlschreiber, Fognini, Melzer, Paire, Querrey, Tomic, Basilashvili.

To do so, we have access to the "head to head" record between them (see Figure 2) in the form of the matrix $R \in \mathbb{R}^{n \times n}$:

$$R_{i,j} = \text{« number of wins of player } i \text{ against player } j \text{ ».} \tag{4}$$

We will use the approach of the previous section to rank the players. In our case, instead of a "drunk surfer" we will consider a "drunk spectator". At time t the value $X_t \in \{1, \ldots, n\}$ indicates which tennis player the spectator believes to be the best. At time $t+1$, the spectator picks uniformly at random a game played by its favorite player X_t against one of the other players, x. If the game was won by X_t, then the spectator still believes that X_t is the best: $X_{t+1} = X_t$. Otherwise the spectator changes his mind: $X_{t+1} = x$.

This can be modeled by a Markov chain with transition matrix:

$$P_{i,j} = \begin{cases} V_j/G_j & \text{if } i = j \\ R_{i,j}/G_j & \text{otherwise,} \end{cases}$$

where V_j denotes the total number of victories of player j and where G_j denotes the total number of game played by j:

$$V_j = \sum_{i=1}^{n} R_{j,i} \qquad \text{and} \qquad G_j = \sum_{i=1}^{n} R_{i,j} + R_{j,i}.$$

Let μ be the "Perron-Frobenius" eigenvector of P. The vector μ is displayed on Figure 1. Applying Corollary 2.2-Theorem 2.2 to the matrix P we get that the "drunk spectator" will (in the $t \to \infty$ limit) spend a fraction μ_i of its time thinking that the player i is the best. The values $(\mu_1, \ldots, \mu_n)$ can therefore be used to rank the players. We obtain the following order (see also Figure 1)

Federer (14.4%), Djokovic (13.7%), Nadal (13.6%), Murray (5.8%), Ferrer (3.0%), Del Potro (2.8%), Berdych (2.5%), Roddick (2.4%), Wawrinka (2.3%), Tsonga (2.1%), Nishikori (1.6%), Nalbandian (1.6%), Hewitt (1.5%), Monfils (1.5%), Davydenko (1.5%), Cilic (1.4%), Soderling (1.4%), Verdasco (1.2%), Gonzalez (1.2%), Raonic (1.2%), Ljubicic (1.2%), Gasquet (1.2%), Simon (1.1%), Thiem (1.1%), Isner (1.0%), Zverev (1.0%), Youzhny (1.0%), Robredo (0.9%), Kohlschreiber (0.9%), Ferrero (0.9%), Stepanek (0.8%), Safin (0.8%), Dimitrov (0.8%), Almagro (0.7%), Baghdatis (0.7%), Blake (0.7%), Anderson (0.7%), Goffin (0.7%), Coria (0.7%), Bautista Agut (0.6%), Monaco (0.6%), Fognini (0.6%), Querrey (0.6%), Melzer (0.6%), Dolgopolov (0.5%), Coric (0.5%), Pouille (0.4%), Tsitsipas (0.4%), Sock (0.4%), Paire (0.3%), Medvedev (0.3%), Khachanov (0.3%), Tomic (0.2%), Basilashvili (0.1%).

References

[1] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

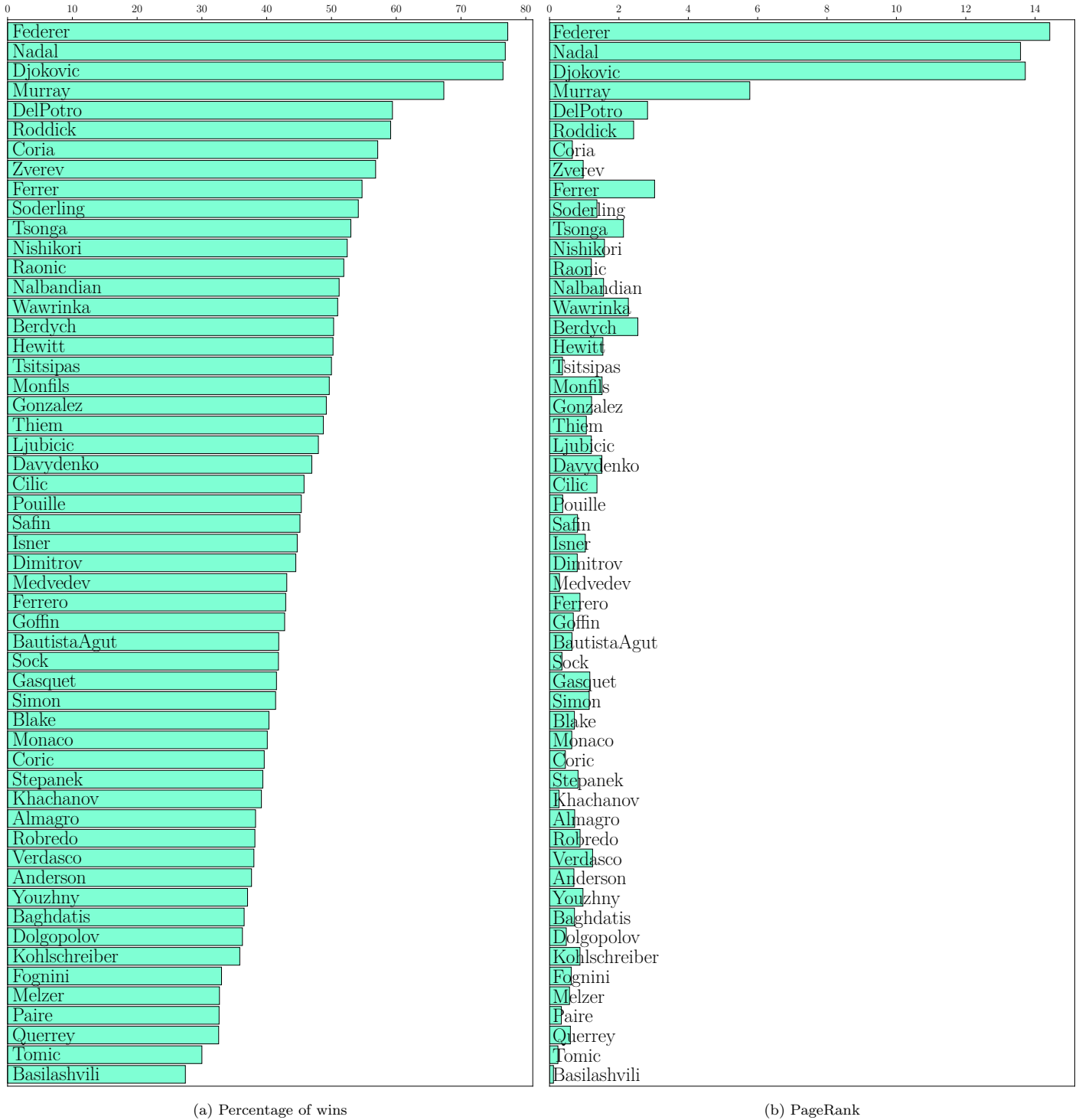[2] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer Science & Business Media, 2006.

(a) Percentage of wins

(b) PageRank

Figure 1: Comparison of the ranking by the percentage of wins (on the left) and the ranking using PageRank.

	Fe	Na	Dj	Mu	De	Ro	Co	Zv	Fe	So	Ts	Ni	Ra	Na	Wa	Be	He	Ts	Mo	Go	Th	Lj	Da	Ci	Po	Sa	Is	Di	Me	Fe	Go	Ba	So	Ga	Si	Bl	Mo	Co	St	Kh	Al	Ro	Ve	An	Yo	Ba	Do	Ko	Fo	Me	Pa	Qu	To	Ba
Fe	0	15	22	14	18	21	3	3	17	16	11	7	11	11	23	20	18	1	10	12	2	13	19	9	1	10	6	7	3	10	7	8	4	18	7	10	4	4	14	1	5	11	7	6	17	7	5	14	4	4	7	3	4	1
Na	24	0	26	17	11	7	4	5	26	6	8	11	7	5	18	20	7	4	14	7	9	7	5	6	2	2	7	12	0	7	4	3	4	16	8	4	7	2	7	6	15	7	17	5	13	9	7	15	11	3	4	4	3	3
Dj	25	28	0	25	16	4	2	3	16	6	17	16	9	4	19	25	6	1	15	1	6	7	6	17	1	0	9	8	3	2	5	7	1	13	11	3	8	3	13	1	5	7	11	8	7	8	6	10	8	3	1	8	6	2
Mu	11	7	11	0	7	8	0	1	14	3	14	9	9	5	11	11	1	0	4	1	2	4	6	12	4	0	8	8	0	3	6	3	0	8	16	2	5	2	7	1	5	6	13	6	4	5	4	5	4	7	2	7	5	0
De	7	6	4	3	0	4	0	2	7	4	5	6	3	1	4	5	2	1	2	2	4	1	4	11	0	1	8	6	0	1	2	3	1	7	5	2	1	1	3	3	4	2	5	7	4	4	5	7	1	5	1	3	2	1
Ro	3	3	5	3	1	0	5	0	4	2	2	1	1	4	1	6	7	0	3	9	0	7	5	1	0	4	4	1	0	5	0	0	1	3	2	9	1	0	7	0	2	11	10	2	3	3	1	4	2	10	0	6	1	0
Co	0	1	2	0	0	0	0	4	1	0	0	0	2	0	2	0	0	0	5	0	3	2	0	0	1	0	0	0	3	0	0	0	1	0	1	2	0	0	0	1	3	2	0	6	0	0	0	0	2	0	0	0	0	0
Zv	3	0	2	0	0	0	0	0	5	0	1	2	1	0	2	2	0	1	0	0	2	0	0	6	0	0	5	2	4	0	2	4	1	4	4	0	1	1	0	2	1	0	1	4	3	1	1	2	3	0	1	0	0	1
Fe	0	6	5	6	6	7	1	3	0	4	3	4	4	9	7	8	3	0	3	5	1	6	2	4	1	1	7	5	0	7	2	3	2	10	8	1	5	0	8	0	15	8	14	3	4	4	10	11	11	7	3	3	4	1
So	1	2	1	2	1	4	0	0	10	0	5	1	0	1	2	7	2	0	3	5	0	2	7	2	0	0	1	0	0	1	0	0	0	2	5	2	4	0	5	0	5	5	5	5	0	5	3	0	1	3	2	0	4	0
Ts	6	4	6	2	2	1	0	2	1	0	0	3	2	1	3	5	4	1	4	1	2	3	4	2	2	1	2	4	1	3	4	2	3	5	9	3	6	1	3	2	6	1	3	3	3	7	3	11	4	6	3	4	3	1
Ni	3	2	2	2	0	0	1	10	0	6	0	5	0	4	5	0	1	4	0	3	0	2	0	1	1	3	0	2	5	2	1	3	4	1	3	1	3	1	1	1	2	2	3	5	5	2	1	5	3	2	3	7	6	3
Ra	3	2	0	3	2	0	0	0	5	2	0	1	3	6	1	0	3	0	2	0	1	1	3	0	1	2	0	0	3	5	8	3	5	2	1	1	3	0	2	6	4	1	3	3	2	2	1	0	4	5	0	0	0	0
Na	8	2	1	2	3	2	2	0	5	6	1	1	0	0	3	4	3	0	1	5	0	4	7	4	0	3	2	1	0	4	0	0	0	7	2	0	3	0	2	0	4	6	0	0	2	2	0	2	2	1	1	0	0	0
Wa	3	3	5	8	3	3	0	0	7	2	5	7	5	6	0	11	2	1	3	0	3	3	1	12	1	3	1	5	0	3	3	1	1	1	4	3	4	3	3	2	6	3	3	5	3	6	2	5	5	2	9	5	1	0
Be	6	4	3	6	4	5	1	4	8	3	8	1	3	1	5	0	3	0	6	3	2	2	3	6	1	1	7	3	0	2	2	4	2	9	8	3	7	3	3	0	9	7	11	12	6	4	4	9	3	5	4	5	5	0
He	9	4	1	0	3	7	2	0	1	3	0	2	1	3	2	0	0	0	2	2	0	1	4	1	0	6	4	1	0	6	0	0	1	2	0	8	1	0	3	0	1	1	0	1	5	3	0	2	0	7	0	3	0	0
Ts	1	1	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0	0	1	0	0	0	0	0	4	1	0	1	0	0	0	1	0	2	0	1	1	2	0	0	0	2	2	0	0	0	0	1
Mo	4	2	0	2	0	5	0	3	3	0	4	1	3	3	3	1	2	1	0	2	0	3	2	4	2	4	7	4	1	0	2	3	1	10	2	3	2	2	6	0	3	2	3	5	1	4	4	14	4	4	0	3	2	1
Go	1	3	2	2	3	3	1	0	5	4	1	0	1	3	5	4	5	0	0	0	0	4	0	1	0	6	1	0	0	3	0	0	0	1	7	6	0	1	0	2	3	3	1	4	1	1	0	0	2	0	2	0	0	0
Th	4	4	3	1	0	0	0	5	1	0	0	2	1	0	1	0	0	3	5	0	0	0	0	0	1	0	1	2	2	0	3	1	3	0	9	0	1	3	1	1	2	0	0	2	2	0	2	1	3	1	2	3	2	0
Lj	3	2	2	3	1	4	2	0	1	3	3	2	0	5	3	3	0	0	4	4	0	0	4	2	0	2	0	0	0	3	0	0	0	3	4	2	0	1	0	3	5	3	0	8	2	1	3	0	0	2	2	0	0	0
Da	2	6	2	4	3	1	1	0	4	4	2	0	0	5	2	9	0	0	2	6	0	4	0	2	0	4	3	0	0	3	1	0	0	2	3	1	4	0	7	0	1	5	7	1	4	2	2	3	4	6	1	1	0	0
Ci	1	2	2	3	2	2	0	1	2	0	6	6	2	2	2	2	6	1	1	1	0	1	0	1	3	0	2	0	8	4	0	2	3	4	0	2	1	1	2	7	1	1	2	4	10	6	5	5	2	5	3	7	4	6
Po	0	1	0	1	1	0	0	0	2	0	2	0	1	0	0	3	0	0	0	2	0	0	0	0	2	0	0	0	0	0	4	3	2	0	4	1	0	1	2	0	1	1	0	1	0	2	2	0	3	0	0	1	0	0
Sa	2	0	2	1	0	3	1	0	1	2	0	0	0	6	1	2	7	0	0	3	0	2	4	1	0	0	1	0	0	6	0	0	0	4	0	2	0	0	1	0	0	4	1	0	3	1	0	0	2	1	0	1	0	0
Is	2	0	2	0	4	2	0	1	2	0	3	1	5	1	3	2	2	2	4	1	1	0	3	3	0	1	0	2	0	0	2	3	5	2	3	3	3	1	2	0	2	3	1	8	2	8	0	4	2	2	1	3	3	0
Di	2	0	0	2	0	0	1	1	0	0	1	3	0	4	3	0	0	0	1	0	2	0	1	2	2	0	1	0	1	0	7	3	2	3	4	0	1	0	1	0	2	2	4	6	1	8	4	2	4	2	1	3	2	1
Me	0	0	1	1	0	0	0	0	0	2	2	2	0	1	0	0	4	1	0	0	0	4	1	0	0	0	0	1	0	2	0	0	0	1	0	0	0	2	0	0	0	0	2	0	1	1	0	1	1	0	1	1	0	0
Fe	3	2	1	0	2	0	3	0	2	1	1	1	0	3	3	0	4	0	3	4	0	3	2	0	0	6	1	0	0	0	0	0	1	2	3	2	0	1	0	2	3	3	1	4	1	1	1	0	4	0	2	0	0	0
Go	1	1	1	0	3	0	0	0	0	0	3	0	2	0	2	1	0	2	3	0	7	0	0	3	1	0	1	1	0	0	0	3	3	1	3	0	0	4	2	4	3	0	3	1	0	2	1	2	0	2	3	2	1	2
Ba	0	0	3	1	2	0	0	2	1	0	3	1	0	0	1	3	0	0	1	3	0	0	2	3	0	0	3	1	1	1	0	2	0	2	2	1	0	0	4	1	3	1	1	1	1	2	2	2	2	3	1	6	3	1
So	0	0	0	0	1	0	0	0	2	2	0	0	2	3	0	0	0	0	0	0	1	0	3	1	0	3	3	0	0	0	2	0	4	1	1	0	1	1	1	1	0	0	2	2	0	1	1	1	1	1	1	1	2	0
Ga	2	0	1	3	1	2	0	0	3	3	4	7	1	0	2	8	0	1	7	0	2	2	6	2	1	2	3	5	1	0	1	1	0	0	8	2	0	1	0	4	2	7	7	6	1	3	2	2	3	6	3	8	0	0
Si	2	1	1	2	3	2	0	0	2	2	3	0	1	1	3	7	4	0	7	0	2	2	5	6	3	0	0	5	1	0	2	5	2	1	0	1	4	2	2	1	3	1	2	1	4	3	1	5	5	4	6	6	3	3
Bl	3	1	3	0	1	2	3	1	0	2	1	0	1	2	0	2	1	0	2	3	0	2	7	2	0	2	0	0	0	1	0	0	0	2	2	0	2	0	4	0	1	4	2	1	2	1	0	1	1	1	1	0	6	0
Mo	0	1	0	2	1	1	0	0	4	1	0	2	2	1	1	0	0	0	5	0	0	1	1	2	1	2	1	3	0	2	1	1	1	1	3	0	0	0	2	0	5	3	7	2	1	1	0	2	3	7	3	5	1	0
Co	2	2	0	2	1	0	0	3	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	2	0	1	0	3	0	0	3	2	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	2	0	0	2	1	1
St	2	0	1	2	3	1	0	3	2	1	0	0	1	4	1	1	0	3	4	0	2	4	3	0	2	3	0	0	3	0	0	0	3	2	3	5	0	0	0	4	2	4	2	4	3	2	6	2	0	1	1	3	0	0
Kh	0	1	0	1	0	0	1	0	1	2	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	4	1	0	1	2	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	2	0	0	0	0	0	2
Al	0	1	0	1	0	0	0	0	1	3	1	1	0	2	3	4	2	0	3	0	0	2	3	1	0	3	2	1	0	4	3	0	2	3	2	1	6	0	1	0	0	8	5	1	1	2	3	7	4	3	2	5	0	0
Ro	1	0	2	2	0	0	1	2	2	1	1	0	0	3	6	4	1	0	3	2	1	1	2	3	0	6	1	3	0	2	1	0	1	3	5	5	3	3	0	5	0	1	0	6	1	3	3	0	5	5	4	1	4	0
Ve	0	3	4	3	1	3	0	2	7	2	2	2	3	3	3	4	0	0	3	3	4	1	2	5	1	0	1	3	1	3	3	4	0	8	3	0	5	0	3	2	7	5	0	5	2	0	3	4	4	6	0	1	1	0
An	0	1	1	2	0	2	0	0	3	0	0	4	1	0	4	0	2	1	1	0	7	0	1	1	1	0	4	2	0	0	0	0	0	2	4	4	0	1	4	1	1	0	0	4	0	1	2	2	4	4	2	1	8	1
Yo	0	4	3	0	0	2	2	0	5	1	4	1	1	2	3	6	2	1	3	1	0	1	2	5	0	0	2	0	1	1	0	0	0	5	8	2	4	0	4	0	5	2	0	2	0	4	3	3	1	5	0	0	2	0
Ba	1	1	0	3	2	1	0	1	2	2	0	4	1	3	0	3	2	0	1	0	1	4	2	1	1	0	1	0	2	4	1	0	3	3	2	1	0	2	0	1	3	3	1	5	0	1	5	1	1	3	2	1	1	0
Do	0	2	0	0	0	0	0	0	4	1	3	1	1	1	2	2	0	0	0	1	1	0	1	1	0	0	0	1	1	1	1	0	1	0	3	2	1	2	1	2	2	1	0	2	1	4	0	1	5	1	0	2	6	0
Ko	0	1	2	1	2	2	0	3	3	4	1	0	1	1	0	2	2	0	2	0	2	0	2	7	1	0	4	0	1	3	1	3	1	2	5	2	2	2	2	3	3	3	6	0	7	3	3	0	7	4	4	2	1	1
Fo	0	4	0	3	1	0	0	1	0	0	2	1	0	0	2	2	0	0	4	1	1	0	1	1	2	1	0	2	1	0	1	7	1	2	0	0	2	2	0	3	5	3	1	3	3	1	2	0	0	1	1	2	0	0
Me	1	1	1	0	1	0	0	1	2	0	0	1	2	1	2	2	0	0	1	2	1	5	1	3	0	4	3	0	0	3	3	2	0	2	2	0	1	0	2	0	3	4	3	1	2	1	1	0	3	0	1	1	0	0
Pa	0	0	1	0	1	0	0	0	1	0	1	2	1	0	3	0	0	1	1	0	0	0	1	1	1	0	0	2	0	1	3	0	0	1	3	1	1	0	0	1	0	4	0	0	2	2	1	3	3	0	0	0	2	3
Qu	0	1	2	2	1	2	0	0	1	0	4	4	2	0	2	1	2	0	0	0	0	1	0	1	1	2	0	5	0	0	0	1	0	0	2	1	1	0	0	2	1	4	8	3	3	4	2	0	2	0	0	0	0	2
To	0	0	0	0	0	0	0	0	2	1	0	2	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	1	1	2	0	1	0	1	1	0	0	0	6	4	0	0	4	2	2	0	0	3	0
Ba	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	2	0	1	0	0	2	0	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0

Figure 2: The matrix R given by (4)

9