

week08-03-recap-markov

March 14, 2024

1 Math087 - Mathematical Modeling

Tufts University – Department of Math
George McNinch george.mcninch@tufts.edu
Fall 2020

1.1 Course material (Week 8): Problem session notebook

2 Introduction

Last week, we introduced some finite-state machines whose transition behavior could be described by a matrix. This week, we investigate properties of such matrices, by studying their eigenvalues and eigenvectors.

3 Eigen-stuff

Recall that a number $\lambda \in \mathbb{R}$ is an *eigenvalue* of $A \in \mathbb{R}^{n \times n}$ if there is a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ for which

$$A\mathbf{v} = \lambda\mathbf{v};$$

\mathbf{v} is then called an *eigenvector*.

If A is diagonal – e.g. if

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$$

– it is easy to see that each standard basis vector \mathbf{e}_i is an eigenvector, with corresponding eigenvalue λ_i (the (i, i) -the entry of A).

4 Diagonalizable matrices

Now suppose that A is an $n \times n$ matrix, that $\mathbf{v}_1, \dots, \mathbf{v}_n$ are eigenvectors for A , and that $\lambda_1, \dots, \lambda_n$ are the corresponding eigenvalues. Write

$$P = [\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n]$$

for the matrix whose columns are the \mathbf{v}_i .

Theorem: If the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent – equivalently, if the matrix P is invertible – then

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} = \text{diag}(\lambda_1, \dots, \lambda_n)$$

i.e. $P^{-1}AP$ is the diagonal matrix $n \times n$ matrix whose diagonal entries are $\lambda_1, \dots, \lambda_n$.

5 Diagonalizable matrices & power iteration

Theorem: Let A be a diagonalizable $n \times n$, with n linearly independent eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. As before, write

$$P = [\mathbf{v}_1 \quad \dots \quad \mathbf{v}_n].$$

a) Suppose $|\lambda_i| < 1$ for all i . Then $A^m \rightarrow \mathbf{0}$ as $m \rightarrow \infty$.

b) Suppose that $\lambda_1 = 1$, and $|\lambda_i| < 1$ for $2 \leq i \leq n$. Any vector $\mathbf{v} \in \mathbb{R}^n$ may be written

$$\mathbf{v} = \sum_{i=1}^n c_i \mathbf{v}_i.$$

If $c_1 \neq 0$, then

$$A^m \mathbf{v} = c_1 \mathbf{v}_1 \quad \text{as } m \rightarrow \infty.$$

If $c_1 = 0$ then

$$A^m \mathbf{v} = \mathbf{0} \quad \text{as } m \rightarrow \infty.$$

5.1 Corollary

Suppose that A is diagonalizable with eigenvalues $\lambda_1, \dots, \lambda_n$, that $\lambda_1 = 1$, and that $|\lambda_i| < 1$ for $i = 2, \dots, n$. Let \mathbf{v}_1 be a 1-eigenvector for A .

Then

$$A^m \rightarrow B \quad \text{as } m \rightarrow \infty$$

for a matrix B with the property that each column of B is either $\mathbf{0}$ or some multiple of \mathbf{v}_1 .

6 Stochastic Matrices

A vector $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_n]^T \in \mathbb{R}^n$ will be said to be a *probability vector* if all of its entries v_i satisfy $v_i \geq 0$ and if

$$[1 \ 1 \ \cdots \ 1] \cdot \mathbf{v} = \sum_{i=1}^n v_i = 1.$$

Let $A = [a_{ij}] \in \mathbb{R}^{n \times n}$. We say that A is a *stochastic matrix* if $a_{ij} \geq 0$ for all i, j and if

$$[1 \ 1 \ \cdots \ 1] \cdot A = [1 \ 1 \ \cdots \ 1];$$

in words, A is a stochastic matrix if each column of A is a probability vector.

Proposition: Let A be a stochastic matrix.

- a) A has an eigenvector with eigenvalue 1.
- b) Let λ be any eigenvalue of a A . Then $|\lambda| \leq 1$.
- c) If \mathbf{w} is an eigenvector of A with eigenvalue λ satisfying $\lambda \neq 1$ then $[1 \ 1 \ \cdots \ 1] \mathbf{w} = 0$.

Corollary: Suppose that the stochastic matrix A is diagonalizable, and that the *1-eigenspace* of A has dimension 1. Let \mathbf{v} be an eigenvector for A with eigenvalue 1, and set $c = [1 \ 1 \ \cdots \ 1] \mathbf{v}$. Then $\mathbf{w} = \frac{\mathbf{v}}{c}$ is a probability vector, and

$$A^m \rightarrow B \quad \text{as } m \rightarrow \infty$$

for a stochastic matrix B . Each column of B is equal to \mathbf{w} .

Theorem: (Perron-Frobenius) Let G be a transition diagram for a Markov chain, and suppose that G is strongly connected and aperiodic. Let P be the corresponding stochastic matrix. The multiplicity of the eigenvalue $\lambda = 1$ for P is 1 – i.e.

$$\dim \text{Null}(P - I_n) = 1.$$

All other eigenvalues λ satisfy $|\lambda| < 1$.

There is a 1-eigenvector \mathbf{v} which is a probability vector.

```
[ ]: import numpy as np

from numpy.random import default_rng

from numpy.linalg import eig, matrix_power

rng = default_rng()

def rand_stoch(n):
    v=np.array([rng.random(n) for i in range(n)])
```

```
f=np.ones(n)@v
return (1/f)*v
```

```
[ ]: T=rand_stoch(8)
      np.ones(8)@T
```

```
[ ]: e_vals,e_vecs = eig(T)
```

```
[ ]: e_vals
```

```
[ ]: w=e_vecs[:,0]
      v = (1/(np.ones(8)@w))*w
      v
```

```
[ ]: vv=(matrix_power(T,200)[: ,0])
```

```
[ ]: vv
```

```
[ ]: v-vv < 1e-7*np.ones(8)
```

```
[ ]: (v-vv < 1e-7*np.ones(8)).all()
```

6.1 How to rank-order the entries in a vector??

Python can sort – but sometimes you don't just want the sorted values.

```
[ ]: vv
```

```
[ ]: ll = [(i,vv[i]) for i in range(8)]
      ll
```

```
[ ]: ll.sort(key=lambda x:(-1)*x[1])
```

```
[ ]: ll
```

```
[ ]: def top_ten(n,it):
      T=rand_stoch(n)
      vv=matrix_power(T,it)[: ,0]
      ll=[(i,vv[i]) for i in range(n)]
      ll.sort(key=lambda x:(-1)*x[1])
      iter_string = "\n".join([f"{l[0]:3d} - {l[1]:.5f}" for l in ll[0:10]])

      e_vals,e_vecs = eig(T)
      w=e_vecs[:,0]
      ww = (1/(np.ones(n)@w))*w
      kl=[(i,ww[i]) for i in range(n)]
      kl.sort(key=lambda x:(-1)*x[1])
      eig_string = "\n".join([f"{l[0]:3d} - {l[1]:.5f}" for l in kl[0:10]])
```

```
return iter_string + "\n\n" + eig_string
```

```
[ ]: print(top_ten(50,2))
```