Stat 230 Summer 2020 I

Matrices

This document contain hands-on computation formulae for linear algebra that we need to answer some of the more interesting questions in discrete time finite state space Markov Chains. **Note that you can solve the homework problems (or any problem on exam) without knowing linear algebra** by e.g. constructing a system of equations. The purpose of this document is just to introduce matrices to make the Markov chains more tangible, and the exposition in Durrett easier to follow. For a deeper understanding you should take a class in linear algebra, which is also fundamentally important for understanding statistics in more general.

A vector $\mathbf{p_j}$ of length n is an ordered list of n real numbers numbers:

$$\mathbf{p_j} := \{p_{1,j}, p_{2,j}, \ldots, p_{n,j}\}.$$

If we have n such vectors, we can create an $n \times n$ table $\mathbf{P}$ where each vector corresponds to a row of the table:

$$\mathbf{P} = \begin{pmatrix} \mathbf{p_1} \\ \mathbf{p_2} \\ \vdots \\ \mathbf{p_n} \end{pmatrix} = \begin{pmatrix} p_{1,1} & p_{2,1} & \cdots & p_{n,1} \\ p_{1,2} & p_{2,2} & \cdots & p_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1,n} & p_{2,n} & \cdots & p_{n,n} \end{pmatrix}$$

This produces an $n \times n$ matrix. The matrices are defined with operations of addition and multiplication. These operations produce another matrix from two matrices (similar to numbers).

The matrix addition is defined as addition of the elements:

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$
$$c_{ij} = a_{ij} + b_{ij}$$

The matrix multiplication is a bit more complicated operation. The $(i, j)th$ element of the product $\mathbf{AB}$ is obtained as a dot product of the ith row of matrix $\mathbf{A}$ and the jth column of matrix $\mathbf{B}$:

$$\mathbf{C} = \mathbf{AB}$$
$$c_{ji} = \sum_{k=1}^{n} a_{ki} b_{jk}.$$

Warning: in general

$$\mathbf{AB} \neq \mathbf{BA}.$$

For our purposes, we need specifically the multiplication in case $\mathbf{A} = \mathbf{B}$. This is called the matrix power, and it is defined recursively $\mathbf{A^n} = \mathbf{AA^{n-1}}$, for $n \geq 2$.

A matrix $\mathbf{P}$ can also multiply[1] a vector $\mathbf{q}$. This produces another vector $\mathbf{r}$:

$$\mathbf{r} = \mathbf{qP},$$

which is defined through its elements

$$\mathbf{r}_i = \sum_{j=1}^{n} \mathbf{P}_{ij}\mathbf{q}_j,$$

i.e. the ith element of $\mathbf{r}$ is the dot product of the ith column of $\mathbf{P}$ and vector $\mathbf{q}$.

The matrix multiplication is *associative*, meaning that e.g.

$$\mathbf{q}\big(\mathbf{P}^2\big) = \mathbf{q}\big(\mathbf{PP}\big) = \big(\mathbf{qP}\big)\mathbf{P}.$$

Matrices in R

Matrices can be defined in R as follows. Note the matrix product in R is "%*%". Do not use *, that gives you the element-wise multiplication.

There are packages for computing matrix powers, but we can also compute them with a for loop.

Below P is a transition matrix of a certain Markov chain, Q is the 100-step transition distribution, and $q\% * \%Q$ is the distribution $\Pr(X_{100}|X_0 = 1)$.

```
# Define 4 vectors of length 4
p1=c(2/3,1/3,0,0)
p2=c(1/2,0,1/2,0)
p3=c(0,1/2,0,1/2)
p4=c(0,0,1/3,2/3)
# Stack the vectors together to form a matrix
P=matrix(c(p1,p2,p3,p4),ncol=length(p1),nrow=length(p1),byrow = TRUE)
# Define another vector
q=c(1,0,0,0)

#Multiply the vector with a matrix
q%*%P

# A loop to construct matrix power
```

[1] In the context of Markov chains, we typically multiply from the right.

```
iter=100
Q=P
for (i in 1:iter) {
  Q=Q%*%P
}
q%*%Q
```