Recitation 10. May 7

Focus: positive definite matrices, Markov matrices.

Definition. A symmetric matrix S is called *positive definite* if all of its eigenvalues are positive. It is *positive semidefinite* all of its eigenvalues are nonnegative, that is we allow zeroes.

Definition. A matrix A is called a *Markov matrix* if all of its entries are nonnegative and the elements in each column sum up to one. It is called a *positive Markov matrix* if in addition we require all matrix entries to be positive.

Fact. A Markov matrix A always has an eigenvalue equal to one, because columns of the matrix $A-I$ lie in the hyperplane $x_1+\cdots+x_n=0$. A nonpositive Markov matrix can have more than one largest eigenvalue, take for example I.

Definition. A *steady state* of a positive Markov matrix A is the unique vector v which is an eigenvector of A with eigenvalue one and whose coordinates sum up to one. It is called "steady vector", because any positive vector x whose coordinates sum up to one converges to v as we iteratively apply A, that is $\lim_n A^n x = v$.

1. Let S be a positive definite matrix. Show that then for any nonzero vector v, we have $v^T S v > 0$.

Solution: We know that a symmetric matrix can be diagonalized: $S = Q\Sigma Q^{-1}$ with an orthogonal Q. Since S is positive definite, all diagonal elements of Σ are positive, call them $\sigma_1, \ldots, \sigma_n$. Let q_i denote the ith column of Q. Then take any vector v and consider its decomposition with respect to the basis of eigenvectors: $v = v_1 q_1 + \cdots + v_n q_n$. Now plug this into the formula, and use e_i to denote the ith standard basis vector:

$$v^T SV = \left(\sum_i v_i q_i^T\right) Q\Sigma Q^T \left(\sum_i v_i q_i\right) = \left(\sum_i v_i e_i^T\right) \Sigma \left(\sum_i v_i e_i\right) =$$
$$= \left(\sum_i v_i e_i^T\right)\left(\sum_i \sigma_i v_i e_i\right) = \sum_i v_i^2 \sigma_i > 0.$$

2. *(Strang, problem 6.5.30.)* The graph of $z = x^2 + y^2$ is a bowl opening upward, or *convex*. The graph of $z = -x^2 - y^2$ is a downward bowl, which means that it is *concave*. The graph of $z = x^2 - y^2$ is a saddle. What is a condition on a, b, c for $z = F(x,y) = ax^2 + 2bxy + cy^2$ to have a saddle point at $(0,0)$?

Solution: Note that we can rewrite the function $F(x,y)$ as follows:

$$F(x,y) = ax^2 + 2bxy + cy^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Now we can take the derivative of this function with respect to $w = \begin{pmatrix} x \\ y \end{pmatrix}$:

$$\mathrm{d}F(x,y) = \mathrm{d}w^T \begin{pmatrix} a & b \\ b & c \end{pmatrix} w + w^T \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mathrm{d}w = 2w^T \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mathrm{d}w.$$

Note that the last transition was possible because the 2×2 matrix is symmetric, and this does not work in general.

So the differential can be written as the following row vector (also called a *covector*):

$$\frac{\mathrm{d}F(x,y)}{\mathrm{d}w} = 2w^T \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Taking the second derivative with respect to w^T will now look like this:

$$\mathrm{d}\frac{\mathrm{d}F(x,y)}{\mathrm{d}w} = 2\mathrm{d}w^T \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Therefore, we get the following formula: $\frac{\mathrm{d}^2 F(x,y)}{\mathrm{d}w^2} = 2\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

We get a saddle point when this matrix is indefinite.

3. If A and B are two Markov matrices, then show that their product AB is Markov as well. Further, derive that then any power A^k, $k > 0$, of a Markov matrix is Markov.

Solution: Write A as a column of rows $\begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}$ and write each row in terms of its elements $a^i = \begin{pmatrix} a^i_1 & \cdots a^i_n \end{pmatrix}$. Write B as a row of its columns: $B = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}$ and use a similar notation b^k_j for the matrix element in row k and column j.

The element in position (i,j) in AB is $(AB)_{ij} = a^i b_j$. And now we want to check the Markov property that $\sum_i (AB)_{ij} = 1$. So let us check it. For the check, note that we can formulate the Markov property of A as follows: when

$$\sum_i (AB)_{ij} = \sum_i a^i b_j = \sum_i \sum_k a^i_k b^k_j = \sum_k \left(\sum_i a^i_k \right) b^k_j = \sum_k 1 \cdot b^k_j = 1.$$

4. *Weather prediciton. (Inspired by en.wikipedia.org/wiki/Examples_of_Markov_chains.)*

   Last May in Boston, there were 10 rainy days and 21 days without precipitation. There were 3 occasions where a rainy day followed a rainy day, 7 occasions where a dry day followed a rainy day. After a dry day, a rainy day followed on 7 occasions and another dry day happened on 13 occasions. (Note than since there are 31 days in May, there are 30 pairs of consecutive days.)

   a) With the first coordinate corresponding to a rainy day and the second – to a dry day, write the vector a_1 of probabilities for what happens after a rainy day and the vector a_2 – for after a dry day.
   b) Using the results from part (a), write the Markov matrix A corresponding to this setup.
   c) Find a steady vector v for A.
   d) Normalize v so that the sum of its coordinates equals to 1 – this will be the steady state.
   e) Compare probability of having a rainy day in May 2018 with the first coordinate in the steady state vector.

Solution:

a) $a_1 = \begin{pmatrix} 30\% \\ 70\% \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.7 \end{pmatrix}$; $a_2 = \begin{pmatrix} \frac{7}{20} \\ \frac{13}{20} \end{pmatrix} = \begin{pmatrix} 0.35 \\ 0.65 \end{pmatrix}$.

b) $A = \begin{pmatrix} a_1 & a_2 \end{pmatrix} = \begin{pmatrix} 0.3 & 0.35 \\ 0.7 & 0.65 \end{pmatrix}$.

c) We know that a steady vector is an eigenvector with eigenvalue one, so it lies in the kernel of the matrix $A - I = \begin{pmatrix} -0.7 & 0.35 \\ 0.7 & -0.35 \end{pmatrix}$, and we can find one vector with this property: $v = \begin{pmatrix} 0.35 & 0.7 \end{pmatrix}$.

d) $v_{st} = \frac{1}{0.35+0.7} \begin{pmatrix} 0.35 & 0.7 \end{pmatrix} = \begin{pmatrix} \frac{0.35}{1.05} \\ \frac{0.7}{1.05} \end{pmatrix} = \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}$.

e) The probability of having a rainy day in May 2018 is $\frac{10}{31} \approx 32\%$, which is very close to the first coordinate in v_{st}, which is $\frac{1}{3} \approx 32\%$. The fact that these quantities are similar can be explained as follows: Markov matrix gives a good simple model of weather prediction. Then, whatever day we start with, applying Markov matrix 30 times to this day will give as something close to v_{st}. So in the long run, we should expect a rainy day in May with probability approximately 33%.

5. If a Markov matrix A has the steady state $(1, \ldots, 1)^T$, then what can you say about the rows of this matrix?

Solution: In each row, the sum of the entries must be equal to one.