CS4642: Homework #3

Due on Friday, October 21, 2013

Hongyuan Zha 3:00pm

Nathan Korzekwa

Contents

"Yet you did not have the wit to see it. Your love of the Halfling's leaf has clearly slowed your mind." – Saruman

Problem 1

(a)

$$A = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}$$

The characteristic polynomial is given by

$$\begin{aligned} det(A - \lambda I) &= (1-\lambda)^2 - 4 \\ &= \lambda^2 - 2\lambda - 3 \end{aligned}$$

(b) and (c)

$$\lambda = -1, 3$$

(d)

To find the eigenspace for $\lambda = -1$, we solve:

$$\begin{aligned} (A + I)\vec{x} &= \vec{0} \\ \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} \vec{x} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \vec{x} &= c \begin{bmatrix} -2 \\ 1 \end{bmatrix} \end{aligned}$$

Equivalently, to find the eigenspace for $\lambda = 3$, we solve:

$$\begin{aligned} (A - 3I)\vec{x} &= \vec{0} \\ \begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix} \vec{x} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \vec{x} &= c \begin{bmatrix} 2 \\ 1 \end{bmatrix} \end{aligned}$$

So, from these eigenspaces, it is easy to see that two eigenvectors for this matrix are $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

(e)

$$\begin{aligned} \vec{x_0} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \vec{x_1} &= \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ \vec{x_1} &= \begin{bmatrix} 5 \\ 2 \end{bmatrix} \end{aligned}$$

(f)

Power iteration will yield the eigenvector associated with the eigenvalue that is largest in magnitude – in this case, $\lambda = 3$, which is:

$$\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(g)

The Raleigh quotient is simply given by $\lambda = \frac{\vec{x}^T A \vec{x}}{\vec{x}^T \vec{x}}$

$$\lambda = \frac{7}{4}$$

(h)

Because in an inverse matrix, the eigenvalues are the reciprocals of those in the original matrix, it will converge to the eigenvector associated with the eigenvalue of smallest magnitude, in this case $\lambda = -1$:

$$\vec{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

(i)

If a matrix A has $\lambda_1, \lambda_2, ..., \lambda_n$ eigenvalues, then matrix $A - \sigma I$ has $\lambda_1 - \sigma, \lambda_2 - \sigma, ..., \lambda_n - \sigma$ eigenvalues. So, in the current problem, the eigenvalues become: $\lambda = -3, 1$. Then, when we invert the matrix, the eigenvalues become $\lambda = -1/3, 1$. So, from this, we can glean that power iteration would converge to the 2nd eigenvector, which is:

$$\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(j)

The matrix A given in the problem is not symmetric (and accordingly has non-zero off-diagonal entries), so QR iteration will yield an upper-triangular matrix.

Problem 2

(a)

In a rank-one matrix, all of the columns lie in the same vector-space, meaning they are all multiples of each other. By the definition of matrix multiplication, it is easy to see that uv^T produces a matrix whose columns are multiples of the same vector.
Observe:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

$$\begin{aligned} \vec{u}\vec{v}^T &= \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \\ &= \begin{bmatrix} v_1u_1 & v_2u_1 & \dots & v_nu_1 \\ v_1u_2 & v_2u_2 & \dots & v_nu_2 \\ \vdots & \vdots & \vdots & \vdots \\ v_1u_n & v_2u_n & \dots & v_nu_n \end{bmatrix} \end{aligned}$$

(b)

The kth row in the matrix uv^T is of the form:

$$\begin{bmatrix} v_1u_k & v_2u_k & \dots & v_nu_k \end{bmatrix}$$

So, a vector multiplication of a row, with say, $\vec{u}$, results in:

$$u_k(v_1u_1 + v_2u_2 + \cdots + v_nu_n)$$

for all k.
This results in a multiple of $\vec{u}$ where $(v_1u_1 + v_2u_2 + \cdots + v_nu_n) = v^Tu = u^Tv$ is the coefficient.

(c)

Because A is a rank-one matrix, all of the other eigenvalues are zero.

(d)

The nature of a rank-one matrix causes the starting vector transformed by A to converge to a multiple of the primary eigenvector in one step, because the columns of the matrix are the eigenvector itself.

Problem 3

$$A = \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.3 \\ 0.1 & 0.1 & 0.6 \end{bmatrix}$$

$$\vec{x}^{(0)} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

(a)

$$\vec{x}^{(3)} = A^3\vec{x}^{(0)}$$

$$\vec{x}^{(3)} = \begin{bmatrix} 0.587x_1 + 0.371x_2 + 0.28x_3 \\ 0.238x_1 + 0.454x_2 + 0.42x_3 \\ 0.175x_1 + 0.175x_2 + 0.3x_3 \end{bmatrix}$$

(b)

$$\vec{x}^{(\infty)} = A^{\infty}\vec{x}^{(0)}$$

$$\vec{x}^{(\infty)} = \begin{bmatrix} 0.45(x_1 + x_2 + x_3) \\ 0.35(x_1 + x_2 + x_3) \\ 0.2(x_1 + x_2 + x_3) \end{bmatrix}$$

$$= \begin{bmatrix} 0.45 \\ 0.35 \\ 0.2 \end{bmatrix}$$

(c)

No, as can be seen above. Since the columns of the matrix are all identical, the matrix-vector multiplication can be represented by factoring out the unitary number in the row, as seen above. And, because there is a constraint on every probability vector $\vec{v}$, such that $v_i > 0$ for all i, and $\|\vec{v}\|_1 = 1$, the vector is independent of its starting state.

(d)

$$A^{\infty} = \begin{bmatrix} 0.45 & 0.45 & 0.45 \\ 0.35 & 0.35 & 0.35 \\ 0.2 & 0.2 & 0.2 \end{bmatrix}$$

The rank of this matrix is obviously one since all of the columns are identical.

Problem 3 continued on next page. . .

(e)

Matrix A can be diagonalized by computing matrices P, and diagonal D, such that $A = PDP^{-1}$. According to the Diagonalization Theorem, P and D are as follows:

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$$

$$P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} = \begin{bmatrix} -0.7448 & -0.7071 & 0.4082 \\ -0.5793 & 0.7071 & -0.8165 \\ -0.3310 & 0 & 0.4082 \end{bmatrix}$$

Another property of the Diagonalization Theorem is this:

$$A^k = PD^kP^{-1}$$

This is very easy to compute since D is diagonal. When we examine D, we notice that D_{11} is 1, but the other diagonal entries are all less than 1. As a result, the matrix D diminishes to:

$$D^\infty = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so

$$A^\infty = PD^\infty P^{-1}$$

Immediately, this explains why the matrix is rank one, since there is only one eigenvalue. Then, with further analysis, we see that D eliminates the all but the first eigenvector in P as well. As a result, the columns of $A^\infty = PD^\infty P^{-1}$ are all a multiple of the first eigenvector of A.

(f)

Yes. Because of the constraints on the matrix, that is, because the columns of a Markov transition matrix A must always sum to 1, $det(A - I) = 0$ is true for all Markov transition matrices. While the columns of matrix A all sum to 1, it follows that the columns of the matrix $A - I$ must all sum to 0. Because of this, the space spanned by $A - I$ is bound by the constraint $\|(A - I)x\|_1 = 0$ for all x. So for $x \epsilon span(A - I)$, $\sum_i x_i = 0$, meaning that one of the components of x is dependent on the others. From this, we lose one degree of freedom, and accordingly $rank(A - I) < n$, where n is the number of columns. Therefore, by the invertible matrix theorem, $det(A - I) = 0$.

(g)

Easy. The set of guaranteed stationary vectors lie in the space spanned by the eigenvector associated with $\lambda = 1$.

(h)

If we know A^∞, then multiplying A^∞ by any vector will result in a stationary vector.

Problem 4

(a)

The code for this problem can be found in "q4.m" in the included zip file. The dominant eigenvector and eigenvalue the method returns is $\lambda_1 = 11.0000, \vec{v}_1 = \begin{bmatrix} 0.5000 \\ 1.0000 \\ 0.7500 \end{bmatrix}$

(b)

The code for this problem can be found in "q4b.m" in the included zip file. After deflating the matrix once to get B, and by transforming the dominant eigenvector for B to be a suitable eigenvector for A, my program calculated:

$$\lambda_2 = -3, \vec{v}_2 = \begin{bmatrix} -0.0000 \\ 0.6839 \\ -1.0258 \end{bmatrix}$$

My program also gets the third eigenvalue/eigenvector pair, but they will not be included in this analysis.

(c)

My program produced:

$$\lambda_1 = 11.0000, \vec{v}_1 = \begin{bmatrix} 0.5000 \\ 1.0000 \\ 0.7500 \end{bmatrix}$$

$$\lambda_2 = -3.0000, \vec{v}_2 = \begin{bmatrix} -0.0000 \\ 0.6839 \\ -1.0258 \end{bmatrix}$$

MATLAB's eig function produced (it was originally in a different order):

$$\lambda_1 = 11.0000, \vec{v}_1 = \begin{bmatrix} 0.3714 \\ 0.7428 \\ 0.5571 \end{bmatrix}$$

$$\lambda_2 = -3.0000, \vec{v}_2 = \begin{bmatrix} -0.0000 \\ -0.5547 \\ 0.8321 \end{bmatrix}$$

The eigenvalues are obviously equivalent, so that checks out. The eigenvectors are not the same, but that's ok as long as they span the same space. A quick point-wise division of each pair reveals them to be ok, since the division reveals a multiple of $\vec{1}$.

Problem 5

(a)

The code for this part of the assignment can be found in "q5.m". I used an ϵ cut-off value on the value of $\|A^{(k)} - A^{(k-1)}\|_F$ to determine convergence. The code works well. For example, for $A = \begin{bmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{bmatrix}$, the code produces:

$$A^{(\infty)} = \begin{bmatrix} 11.0000 & 0.7220 & -6.1864 \\ 0 & -3.0000 & -3.8996 \\ 0 & 0 & -2.0000 \end{bmatrix}$$

Which is consistent with the previous question!

Problem 6

(a)

We have shown previously that for a rank-one matrix $A = \vec{u}\vec{v}^T$, $\vec{u}^T\vec{v}$ is an eigenvalue of A. We can also show that $det(A)$ is the product of the eigenvalues since for all eigenvalues λ of A, $det(A - \lambda I) = 0$. Since $A = \vec{u}\vec{v}^T$ is a rank-one matrix, it's only eigenvalue is $\vec{u}^T\vec{v}$. So $det(A) = \vec{u}^T\vec{v}$. So, from this, $I + \vec{u}\vec{v}^T$ is like shifting A by negative 1, that is, adding 1 to all of its eigenvalues. And since A has only one eigenvalue, the product of its eigenvalues is also incremented by 1, giving us $det(A) = \vec{u}^T\vec{v} + 1$.