Homework 3 (error and iterative methods, markov chains)

Hanzhang Yin

September 27, 2024

Recall that a Markov matrix is an $n \times n$ matrix whose *columns* represent the probabilities of transitioning between n states: the columns sum to 1 and entry M_{ij} is the probability of transitioning from state j to state i. Some sources use the transpose of this matrix.

1. Apply iterative methods (our generalized Jacobi method) to estimate a solution to

$$\begin{bmatrix} 4 & 5 \\ 3 & 5 \end{bmatrix} x = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

a) Pick any splitting matrix besides A^{-1}, and any initial guess besides the actual solution, and determine the result after two iterations.

b) Determine the quantity $\delta = ||I - Q^{-1}A||$ (notation from class/book; Q is the splitting matrix).

c) Determine the actual solution by inverting the matrix.

d) Compare the actual solution to your approximate solution from (a) using the ∞ norm. Then, compare to the error estimate theorem from class. What do you notice?

Solution.
(a):
We pick splitting matrix $Q = \begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix}$, as the diagonal part of A. The Remainder matrix R will be:

$$R = Q - A = \begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix} - \begin{bmatrix} 4 & 5 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 0 & -5 \\ -3 & 0 \end{bmatrix}$$

Selectg initial guess $x^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.
Using Cramer's rule the inverse of Q is:

$$Q^{-1} = \frac{1}{det(Q)} Q = \frac{1}{20} \begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \end{bmatrix}$$

Denote matrix $M = Q^{-1}R$, we can get:

$$M = Q^{-1}R = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 0 & -5 \\ -3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{5}{4} \\ -\frac{3}{5} & 0 \end{bmatrix}$$

Denote vector $c = Q^{-1}b$, we can get:

$$c = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{3}{5} \end{bmatrix}$$

Now, we can denote the first iteration (k = 0) as:

$$x^{(1)} = Mx^{(0)} + c = \begin{bmatrix} \frac{1}{2} \\ \frac{3}{5} \end{bmatrix}$$

Then second iteration (K = 1):

$$x^{(1)} = Mx^{(1)} + c = \begin{bmatrix} 0 & -\frac{5}{4} \\ -\frac{3}{5} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{3}{5} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ \frac{3}{5} \end{bmatrix}$$

$$= \begin{bmatrix} (-\frac{5}{4}) \cdot \frac{3}{5} \\ (-\frac{3}{5}) \cdot \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ \frac{3}{5} \end{bmatrix} = \begin{bmatrix} -\frac{3}{4} + \frac{1}{2} \\ -\frac{3}{10} + \frac{3}{5} \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} \\ \frac{3}{10} \end{bmatrix}$$

(b):
Compute $Q^{-1}A$:

$$Q^{-1}A = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 4 & 5 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & \frac{5}{4} \\ \frac{3}{5} & 1 \end{bmatrix}$$

Then $I - Q^{-1}A$ will be:

$$I - Q^{-1}A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & \frac{5}{4} \\ \frac{3}{5} & 1 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{5}{4} \\ -\frac{3}{5} & 0 \end{bmatrix}$$

Using Infinity Norm to compute δ:

$$\delta = ||I - Q^{-1}A||_\infty = \max\left(|0| + |-\frac{5}{4}|, |-\frac{3}{5}| + |0|\right) = \max\left(\frac{5}{4}, \frac{3}{5}\right) = \frac{5}{4}$$

(c):

$$\det(A) = (4)(5) - (3)(5) = 20 - 15 = 5$$

$$A^{-1} = \frac{1}{5} \begin{bmatrix} 5 & -5 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -\frac{3}{5} & \frac{4}{5} \end{bmatrix}$$

Then the exact solution x will be:

$$x = A^{-1}b = \begin{bmatrix} 1 & -1 \\ -\frac{3}{5} & \frac{4}{5} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1.2 \end{bmatrix}$$

(d): The error after two iteration can be calculated as:

$$e^{(2)} = x - x^{(2)} = \begin{bmatrix} -1 \\ 1.2 \end{bmatrix} - \begin{bmatrix} -\frac{1}{4} \\ \frac{3}{10} \end{bmatrix} = \begin{bmatrix} -\frac{3}{4} \\ 0.9 \end{bmatrix}$$

Then,

$$||e^{(2)}||_\infty = \max\left(|-\frac{3}{4}|, |0.9|\right) = 0.9$$

The error estimate theorem states:

$$||e^{(k)}|| \leq \delta^k ||e^{(0)}||$$

In here, $\delta = \frac{5}{4}$ and $||e^{(0)}||_\infty = \max(|-1|, |1.2|) = 1.2$
The theoretical error bound after 2 iter. hence will be:

$$\delta^k ||e^{(0)}|| = \left(\frac{5}{4}\right) \times 1.2 = \frac{25}{16} \times 1.2 \approx 1.875$$

Noting that actual error: $||e^{(2)}||_\infty = 0.9$, theoretical error bound: $||e^{(2)}||_\infty \leq 1.875$.
Although the error bound suggests divergence for $\delta > 1$, the error decreases initially before increasing as iterations progress, confirming divergence as predicted and consistent with the error estimate theorem.

2. In our steady-state calculation for a Markov matrix M, we determined that all Markov matrices have 1 as an eigenvalue by iterative methods. Iterative methods requires some technical assumptions that we did not discuss. This problem walks you through verifying this fact without without them.

a) Show that a Markov matrix M has 1 as a left eigenvalue (i.e. an eigenvalue of M^T). Hint: the sum of the rows is 1 in a Markov matrix – what left vector multiplication would produce such a row sum? Is it an eigenvector?

b) Show that A and A^T have the same minimal polynomial. Hint: check that $p(A)^T = p(A^T)$ for any polynomial p.

c) Combine the previous two facts to conclude that A has an eigenvector with eigenvalue 1.

Solution.
(a):

Proof. let v be a $n \times 1$ vector s.t. $v = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$, by the hint, since Markov matrix M is row-stochastic (i.e. rows sum to 1), we can have:

$$(M \cdot v)_i = \sum_{j=1}^{n} M_{ij} \cdot v = \sum_{j=1}^{n} M_{ij} = 1$$

Therefore,

$$Mv = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = v$$

Then we can transpose to both sides:

$$(Mv)^T = v^T \Rightarrow v^T M^T = v^T$$

Hence, by definition, v^T is a left eigenvector of M^T corresponding to eigenvalue $\lambda = 1$. □

(b):
by the given hint, we will show that for any matrix A, $p(A)^T = p(A^T)$ holds and analyze from here.

Proof. Note, for any positive integer k, by applying property $(AB)^T = B^T A^T$ recusively, $(A^k)^T = (A^T)^k$ holds.
Let $p(x)$ by any polynomial, we can denote it as $p(x) = \sum_{i=0}^{n} c_i x^i$, then:

$$p(A)^T = \left(\sum_{i=0}^{n} c_i A^i \right)^T = \sum_{i=0}^{n} c_i (A^i)^T = \sum_{i=0}^{n} c_i (A^T)^i = p(A^T)$$

Let $m_A(x)$ be the minial polynomial of A, so $M_A(A) = 0$, then:

$$m_A(A^T) = m_A(A)^T = 0^T = 0$$

Similarly, if $M_{A^T}(A)$ is the minimal polynomial for A^T, then:

$$m_{A^T}(A) = m_{A^T}(A^T)^T = 0^T = 0$$

Noting that, $m_A(x)$ annihilates A and $m_{A^T}(A)$ annihilates A^T, and minial polynomials are unique monic polynomials of least degree:

$$m_{A^T}(x) = m_A(x)$$

□

(c):

Proof. From part (a), denote $M = A$, we get that A^T has an eigenvalue $\lambda = 1$.
Since A^T has an eigenvalue 1, its minimal polynomial $m_{A^T}(x)$ contains factor $(x-1)$
And, from part (b), we know $m_{A^T}(x) = m_A(x)$. Therefore $m_A(x)$ must aslo contains factor $(x-1)$. This means A has an eigenvalue $\lambda = 1$
By definition of eigenvalue, there exists a non-zsero vector v s.t. $Av = \lambda v = v$.
Therefore, A has eigenvalue 1, with corresponding eigenvector. $\square$

3. Recall that we assumed a matrix has a full-rank eigenspace and a unique largest eigenvalue in order to locate its largest eigenvalue (and associated eigenvector) by iterative methods.
We saw above that every Markov matrix M has 1 as an eigenvalue.

1. Prove every eigenvalue of M has norm at most 1. Hint: use (2b) and the ∞ norm, or, equivalently, the Gershgorin circle theorem.
2. Suppose that M is 2×2 with full rank eigenspace, but 1 is a repeated eigenvalue. What does this mean for the original matrix?
3. Suppose that M is 2×2 and has 1 as its only eigenvalue, but its eigenspace is not full rank. What can you say about this situation? Hint: use the ∞-norm and examine the left hand side of $|Mv| \leq ||M|||v| = |v|$ more carefully.

Solution.
(a):

Proof. Let M by an $n \times n$ Markov matrix. The ∞-norm of M is:

$$||M||_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |M_{ij}| = 1$$

By the submultiplicative property of matrix norms and M is non-negative matrix,

$$\rho(M) \leq ||M||_\infty = 1$$

where $\rho(M)$ denotes the spectral radius of M. Hence eveyr eigenvalue λ of M suffices $|\lambda| \leq 1$. $\square$

(b):

Proof. Let M be a 2×2 Markov matrix with eigenvalue $\lambda = 1$ of both algebraic multiplicity and geometric multiplicity 2 (since M is full rank). When algebraic multiplicity = geometric multiplicity, M is diagonalizable. Hence,

$$M = PDP^{-1} = I$$

Where D is the diagonal matrix with egienvlaues on its diagonal, and P is eigenvectors. Given both eigenvalues are 1, $D = I$, the identity matrix, hence:

$$M = P \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = I$$

If a 2×2 Markov matrix M has a full rank eigenspace with 1 as a repeated eigenvlaue, then M must be the identity matrix. $\square$

(c):

Proof. Let M be a 2×2 Markov matrix with eigenvalue $\lambda = 1$ of both algebraic multiplicity 2 but geometric multiplicity 1 (since M is NOT full rank). Then, M is defective and cannot be diagonalized. Its Jordan form by definition, will be:

$$J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

so, $M = PJP^{-1}$ for some invertible P matrix.
Now consider the iterative behavior of M:

$$M^k = PJ^kP^{-1} = P \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} P^{-1}$$

As $k \to \infty$, the off-diagonal entry k causes M^k to grow without bound.
Then, we can compute the ∞-norm of M:

$$||M||_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |M_{ij}| = 1$$

by the hint, for any vector $v \in \mathbb{R}^2$, supposingly:

$$|Mv|_\infty \leq ||M||_\infty |v|_\infty = |v|_\infty$$

But, from the iterative behavior:

$$|Mv|_\infty \leq |v|_\infty$$

yet M^k grows without bound unless v is a eigenvector.
Now, let v be a generalized eigenvector s.t. $(M - I)v = w$, here w is an eigenvector. Then:

$$Mv = v + w$$

Applying the norm:

$$|Mv|_\infty = |v + w|_\infty \leq |v|_\infty$$

Since w is non-zero and not a scalar multiple of v (since M is defective), adding w to v violating the inequality.
Hence, we derived a contradiction.
By proof by contradiction, no such defective 2×2 Markov matrix M exists. □

4. (bonus) Generalize (3c) to show that a Markov matrix with no zeros has a *unique* eigenvector whose associated eigenvalue has norm 1.

Solution.

Proof. Let M be an $n \times n$ Markov matrix with $M_{ij} > 0$ for all i, j. Each row sums to 1:

$$\sum_{j=1}^{n} M_{ij} = 1, \forall i$$

By the Gershgorin Circle Theorem, every eigenvalue λ of M satisties:

$$|\lambda - M_{ii}| \leq \sum_{j \neq i} |M_{ij}|$$

Because $\sum_{j=1}^{n} M_{ij} = 1$, each row sum implies $M_{ii} + \sum_{j \neq i} |M_{ij}| = 1$. Hence,

$$|\lambda| \leq 1$$

The eigenvalue $\lambda = 1$ always exists for a Markov matrix because $M \cdot 1 = 1$, where 1 is the vector with all entires $= 1$.

To show $\lambda = 1$ is the only eigenvalue, we can consider the **subdominant eigenvalues** of M. If M has no zero entries, then M is irreducible and aperiodic, which implies that all other eigenvalues have strictly smaller modulus (i.e., $|\lambda| < 1$ for $\lambda \neq 1$).
Since $\lambda = 1$ is the only eigenvalue with modulus 1 and with multiplicity 1, it has a unique correspdoning eigenvector, which represents the unique stationary distribution of Markov Chain.
Thus, for Markov matrix M with no entries, the only eigenvalue with modulus 1 is $\lambda = 1$, and it has a unique corresponding eigenvector.

$\square$

5. (ungraded bonus) If you know some graph theory, discuss the implications for a Markov matrix whose state diagram, like the weather one we drew in class, is (directed) connected and such that each state has a nonzero probability of remaining the same. Hint: could the fact above apply to a large power of M?

Solution.