CX 4640 Homework 3

Wenqi He

October 15, 2017

3.5

(a)

The vector is obviously within $span(A)$ and therefore cannot be the residue.

(b)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \neq 0$$

The vector is not orthogonal to $span(A)$, so it cannot be the residue.

(c)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} = 0$$

The vector is orthogonal to $span(A)$, therefore it is a possible value for r

3.28

(a)

For all $i \neq j$,

$$P_i P_j = q_i q_i^T q_j q_j^T = q_i (q_i^T q_j) q_j^T = 0$$

Base case:

$$(I - P_2)(I - P_1) = I - P_1 - P_2 + P_2 P_1 = I - P_2 - P_1$$

Inductive Step:
Suppose $(I - P_n)(I - P_{n-1}) \cdots (I - P_1) = I - P_n - \cdots - P_1$, then

$$\begin{aligned}
&(I - P_{n+1})(I - P_n)(I - P_{n-1}) \cdots (I - P_1) \\
&= (I - P_{n+1})(I - P_n - \cdots - P_1) \\
&= (I - P_n - \cdots - P_1) - P_{n+1}(I - P_n - \cdots - P_1) \\
&= I - P_{n+1} - P_n - \cdots - P_1 + \sum_{i=1}^{n} (P_{n+1} P_i) \\
&= I - P_{n+1} - P_n - \cdots - P_1
\end{aligned}$$

(b)

In the classical Gram-Schmidt procedure, during the i-th iteration, $P_i a_k$ is subtracted away, therefore the process is equivalent to

$$\begin{aligned}
q_k &= v_{k-2} - P_{k-1} a_k \\
&= v_{k-3} - P_{k-2} a_k - P_{k-1} a_k \\
&\cdots \\
&= a_k - P_1 a_k - \cdots - P_{k-1} a_k \\
&= (I - (P_1 + \cdots + P_{k-1})) a_k
\end{aligned}$$

(c)

In the modified Gram-Schmidt procedure, during the i-th iteration, $P_i v_{i-1}$ is subtracted away, where v_{i-1} is the intermediate result from last iteration, therefore the process is equivalent to

$$\begin{aligned}
q_k &= (I - P_{k-1}) v_{k-2} \\
&= (1 - P_{k-1})(1 - P_{k-2}) v_{k-3} \\
&\cdots \\
&= (I - P_{k-1}) \cdots (1 - P_1) a_k
\end{aligned}$$

(d)

It's already shown in (a) that (b) and (c) are equivalent. As for (d), take $m = 2$

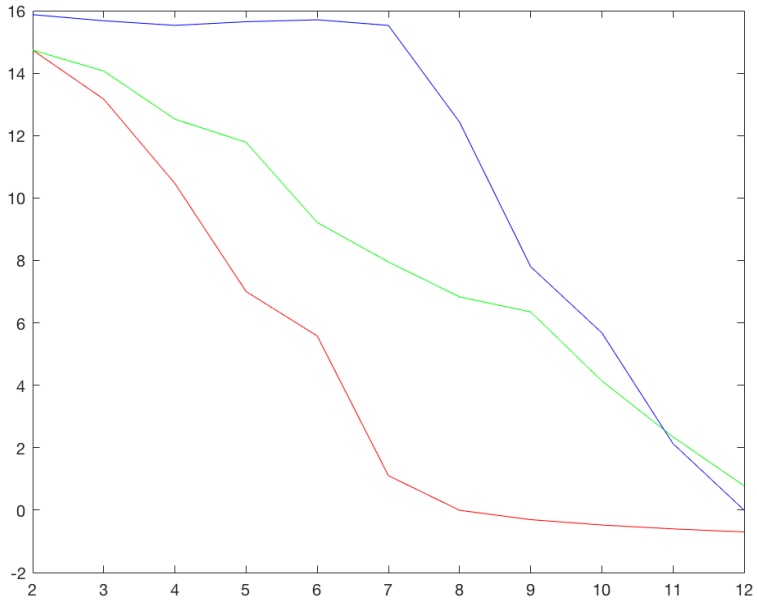
$$\begin{aligned}
&(I - (P_1 + \cdots + P_{k-1}))^2 \\
&= \left(I - \sum_{i=1}^{k-1} P_i\right)\left(I - \sum_{i=1}^{k-1} P_i\right) \\
&= I - \sum_{i=1}^{k-1} P_i - \left(\sum_{i=1}^{k-1} P_i - \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} P_i P_j\right) \\
&= I - \sum_{i=1}^{k-1} P_i - \left(\sum_{i=1}^{k-1} P_i - \sum_{i=1}^{k-1} P_i^2\right)
\end{aligned}$$

Since $P_i^2 = q_i q_i^T q_i q_i^T = q_i (q_i^T q_i) q_i^T = q_i q_i^T = P_i$, the last term evaluates to zero

$$(I - (P_1 + \cdots + P_{k-1}))^2 = I - \sum_{i=1}^{k-1} P_i = I - (P_1 + \cdots + P_{k-1})$$

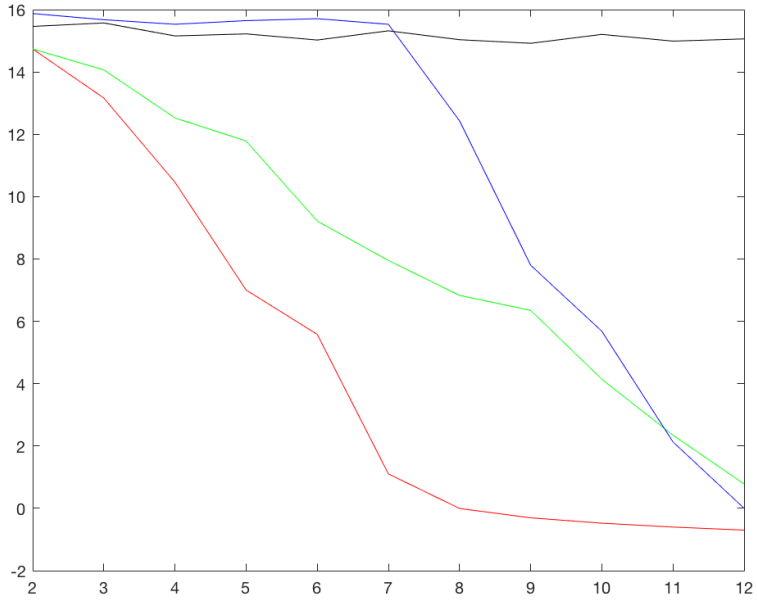
Therefore all three methods are equivalent.

3.12



All three methods become less and less accurate as n grows. Classical Gram-

Schmidt is the least accurate, but applying it twice makes it more accurate than modified Gram-Schmidt, although it takes twice as much time.



In contrast with the three variations of Gram-Schmidt procedure, the Householder method remains accurate as n grows.

4.3

(a)

$$(1-\lambda)^2 - 4$$

(b)

$$\lambda_1 = -1, \quad \lambda_2 = 3$$

(c)

Same as (b)

(d)

For $\lambda_1 = -1$

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} x = 0, \quad x_{\lambda_1} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

For $\lambda_2 = 3$

$$\begin{bmatrix} -2 & 4 \\ 1 & -2 \end{bmatrix} x = 0, \quad x_{\lambda_2} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(e)

$$Ax_0 = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

(f)

It will converge to $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, the eigenvector corresponding to the dominant eigenvalue $\lambda = 3$.

(g)

$$\lambda \approx \frac{x^T A x}{x^T x} = \frac{\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \frac{7}{2} = 3.5$$

(h)

The eigenvalues of A^{-1} are

$$\lambda_1' = \frac{1}{\lambda_1} = -1, \quad \lambda_2' = \frac{1}{\lambda_2} = 1/3$$

The inverse iteration would converge to the eigenvector corresponding to the dominant eigenvalue λ_1', which is

$$x_{\lambda_1'} = x_{\lambda_1} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

(i)

The eigenvalues of $A - \sigma I$ are

$$\lambda_1' = \lambda_1 - \sigma = -3, \quad \lambda_2' = \lambda_2 - \sigma = 1$$

The eigenvalues of $(A - \sigma I)^{-1}$ are

$$\lambda_1'' = \frac{1}{\lambda_1'} = -\frac{1}{3}, \quad \lambda_2'' = \frac{1}{\lambda_2''} = 1$$

The dominant eigenvalue of $(A - \sigma I)^{-1}$, λ_2'', would be obtained from the shifted inverse iteration, which translates to

$$\lambda_2 = 3$$

of the original matrix A

(j)

Since A is not symmetric, it would just converge to triangular form.

4.24

(a)

Since the matrix is of rank one, the columns must be multiples of some real vector u, that is:

$$A = \begin{bmatrix} u \cdot v_1 & u \cdot v_2 & \cdots & u \cdot v_n \end{bmatrix} = u \begin{bmatrix} v_1 & v_2 & \cdots, v_n \end{bmatrix} = uv^T$$

(b)

Multiply A by u on the right:

$$Au = (uv^T)u = u(v^T u) = u(u^T v)$$

Therefore, $u^T v$ is an eigenvalue corresponding to eigenvector u

(c)

The other eigenvalue is 0.
proof: By rank-nullity theorem, A has nullity of $n-1$. For any vector x in the null space, $Ax = 0 = 0x$, which means that 0 is an eigenvalue with multiplicity $n-1$. By spectral theorem, there are no more eigenvalues.

(d)

It only takes 1 iteration.
proof: Any vector can be expressed as a linear combination of the eigenvectors:

$$x = \sum_i c_i x_i = cu + \sum_{x_i \in Null(A)} c_i x_i,$$

After the first iteration:

$$\begin{aligned} Ax &= A\left(cu + \sum_{x_i \in Null(A)} c_i x_i\right) \\ &= A(cu) + 0 = (c\lambda)u \end{aligned}$$

4.12

(a)

$$x^{(3)} = A^3 x^{(0)} = \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.3 \\ 0.1 & 0.1 & 0.6 \end{bmatrix}^3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.587 \\ 0.238 \\ 0.175 \end{bmatrix}$$

(b)

The long-term value must satisfies

$$Ax^{(\infty)} = x^{(\infty)}$$

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

$$\begin{bmatrix} -0.2 & 0.2 & 0.1 \\ 0.1 & -0.3 & 0.3 \\ 0.1 & 0.1 & -0.4 \end{bmatrix} x^{(\infty)} = 0$$

$$x_1^{(\infty)} = 2.25x_3, \quad x_2^{(\infty)} = 1.75x_3$$

$$x^{(\infty)} = \frac{1}{2.25 + 1.75 + 1} \begin{bmatrix} 2.25 \\ 1.75 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.35 \\ 0.2 \end{bmatrix}$$

(c)

No. The equation above has only one solution, therefore all transitions will converge to the same vector regardless of the initial condition.

(d)

Using MATALB to evaluate A^k when k is large:

$$A^{100} = \begin{bmatrix} 0.45 & 0.45 & 0.45 \\ 0.35 & 0.35 & 0.35 \\ 0.2 & 0.2 & 0.2 \end{bmatrix}, \quad A^{200} = \begin{bmatrix} 0.45 & 0.45 & 0.45 \\ 0.35 & 0.35 & 0.35 \\ 0.2 & 0.2 & 0.2 \end{bmatrix}$$

From the result, it can be inferred that

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

The rank of this matrix is 1.

(e)

The eigenvalues and corresponding eigenvectors (unnormalized) of A are:

$$\lambda_1 = 1, \quad v_1 = \begin{bmatrix} -0.7448 \\ -0.5793 \\ -0.3310 \end{bmatrix}$$

$$\lambda_2 = 0.6, \quad v_2 = \begin{bmatrix} -0.7071 \\ 0.7071 \\ 0 \end{bmatrix}$$

$$\lambda_3 = 0.5, \quad v_3 = \begin{bmatrix} 0.4082 \\ -0.8165 \\ 0.4082 \end{bmatrix}$$

The transitions are equivalent to applying power iteration to the initial vector, which converges to to the normalized eigenvector corresponding to the dominant eigenvalue $\lambda_1 = 1$:

$$v_1 = \frac{1}{-0.7448 - 0.5793 - 0.3310} \begin{bmatrix} -0.7448 \\ -0.5793 \\ -0.3310 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.35 \\ 0.2 \end{bmatrix}$$

Therefore, the long-term effect of the Markov chain is just converting any initial vectors to v_1, that is:

$$\lim_{k \to \infty} A^k = \begin{bmatrix} v_1 & v_1 & v_1 \end{bmatrix}$$

(f)

No. If the Markov chain doesn't have a steady-state, that is, if the equation in (b) has no solution, then 1 is not a eigenvalue of A.

(g)

A probability distribution vector that satisfies $Ax = x$ is simply a normalized eigenvector corresponding to eigenvalue 1 .

(h)

By directly solving $(A - I)x = 0$ and then normalize the result.