MA1522: Linear Algebra for Computing

Tutorial 9

Revision

Eigenvalue, Eigenvector, Eigenspace

- Let $\mathbf{A}$ be a square matrix of order n. $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A}$ and $\mathbf{v}$ an associated an eigenvector if $\mathbf{v} \neq \mathbf{0}$ with $\mathbf{A}\mathbf{v} = \lambda \mathbf{v}$.
- The characteristic polynomial of $\mathbf{A}$, is the degree n polynomial $\det(x\mathbf{I} - \mathbf{A})$.
- $\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A}$ if and only if λ is a root of the characteristic polynomial $\det(x\mathbf{I} - \mathbf{A})$.
- The algebraic multiplicity of λ is the largest integer r_λ such that $\det(x\mathbf{I} - \mathbf{A}) = (x - \lambda)^{r_\lambda} p(x)$.
- The eigenvaules of a triangular matrix are the diagonal entries.
- The nonzero (nontrivial) solutions to the homogeneous system $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ are the eigenvector of $\mathbf{A}$ associated to λ.
- The eigenspace associated to λ is $E_\lambda = \left\{ \ \mathbf{v} \in \mathbb{R}^n \ \middle| \ \mathbf{A}\mathbf{v} = \lambda\mathbf{v} \ \right\} = \mathrm{Null}(\lambda\mathbf{I} - \mathbf{A})$.
- The geometric multiplicity of an eigenvalue λ is the dimension of its associated eigenspace, $\dim(E_\lambda) = \mathrm{nullity}(\lambda\mathbf{I} - \mathbf{A})$.

Diagonalization

$\mathbf{A}$ is diagonalizable if there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$ is a diagonal matrix, or $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$.

Theorem

$\mathbf{A}$ is diagonalizable if and only if the characteristic polynomial of $\mathbf{A}$ splits into linear factors,

$$\det(x\mathbf{I} - \mathbf{A}) = (x - \lambda_1)^{r_{\lambda_1}}(x - \lambda_2)^{r_{\lambda_2}} \cdots (x - \lambda_k)^{r_{\lambda_k}},$$

and the geometric multiplicity is equal to the algebraic multiplicity for each eigenvalue λ_i, $\dim(E_{\lambda_i}) = r_{\lambda_i}$.

Algorithm to Diagonalization

1. Compute the characteristic polynomial $\det(x\mathbf{I} - \mathbf{A})$ and find its roots.
2. For each eigenvalue λ_i of $\mathbf{A}$, $i = 1, ..., k$, find a basis S_{λ_i} for the solution space of $(\lambda_i\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$.
3. Let $S = \bigcup_{i=1}^{k} S_{\lambda_i}$. Then $S = \{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n\}$ is a basis for $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A}$.
4. Let $\mathbf{P} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{pmatrix}$, and $\mathbf{D} = \text{diag}(\mu_1, \mu_2, ..., \mu_n)$, where μ_i is the eigenvalue associated to $\mathbf{u}_i$, $i = 1, ..., n$, $\mathbf{A}\mathbf{u}_i = \mu_i\mathbf{u}_i$. Then
$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}.$$

Orthogonally Diagonalization

$\mathbf{A}$ is orthogonally diagonalizable if $\mathbf{A} = \mathbf{PDP}^T$ for some orthogonal matrix $\mathbf{P}$ and diagonal matrix $\mathbf{D}$. $\Leftrightarrow$ $\mathbf{A}$ is symmetric.

Algorithm to Orthogonally Diagonalization

Same as diagonalization, except to apply change the basis for each eigenspace to an orthonormal basis (i.e. apply Gram-Schmidt process to the basis in each eigenspaces).

Tutorial 9 Solutions

Question 1(a)

A father wishes to distribute an amount of money among his three sons Jack, Jim, and John. He wish to distribute such that the following conditions are all satisfied.

(i) The amount Jack receives plus twice the amount Jim receives is \$300.

(ii) The amount Jim receives plus the amount John receives is \$300.

(iii) Jack receives \$300 more than twice of what John receives.

Is it possible for the following conditions to all be satisfied?

Let x, y, z be the amount of money that Jack, Jim, and John receives, respectively. The conditions are

$$\left\{\begin{array}{ccccccc} x & + & 2y & & & = & 300 \\ & & y & + & z & = & 300 \\ x & & & - & 2z & = & 300 \end{array}\right.$$

```
>> A=[1 2 0;0 1 1;1 0 -2]; b=[300;300;300]; rref([A b])
```

This system is inconsistent. So, there are no solution to the distribution problem.

Question 1(b)

If it is not possible, find a least square solution. (Make sure that your least square solution is feasible. For example, one cannot give a negative amount of money to anybody.)

```
>> rref([A'*A A'*b])
```

The least square solutions to the system in (a) is

$$x = 200 + 2t, \quad y = 100 - t, \quad z = t, \quad t \in \mathbb{R}.$$

Need $x, y, z \geq 0$, so $0 \leq t \leq 100$.

Question 2(a)

Suppose $\mathbf{A}$ is a $m \times n$ matrix where $m > n$. Let $\mathbf{A} = \mathbf{QR}$ be a QR factorization of $\mathbf{A}$. Explain how you might use this to write

$$\mathbf{A} = \mathbf{Q}'\mathbf{R}',$$

where $\mathbf{Q}'$ is an $m \times m$ orthogonal matrix, and $\mathbf{R}'$ a $m \times n$ matrix with $m - n$ zero rows at the bottom. This is known as the *full* QR factorization of $\mathbf{A}$.

Write $\mathbf{Q} = \begin{pmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{pmatrix}$ and $\mathbf{R} = \begin{pmatrix} r_{11} & * & \cdots & * \\ 0 & r_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{pmatrix}$. Then

$$\mathbf{A} = \mathbf{QR} = \left(\mathbf{Q}\begin{pmatrix} r_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \mathbf{Q}\begin{pmatrix} r_{12} \\ r_{22} \\ \vdots \\ 0 \end{pmatrix} \quad \cdots \quad \mathbf{Q}\begin{pmatrix} r_{1n} \\ r_{2n} \\ \vdots \\ r_{nn} \end{pmatrix} \right) = \begin{pmatrix} r_{11}\mathbf{q}_1 & r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 & \cdots & r_{1n}\mathbf{q}_1 + \cdots + r_{nn}\mathbf{q}_n \end{pmatrix}.$$

Question 2(a)

Suppose $\mathbf{R}' = \begin{pmatrix} r_{11} & * & \cdots & * \\ 0 & r_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$, then for any $\{\mathbf{q}_{n+1}, ..., \mathbf{q}_m\}$, if $\mathbf{Q}' = \begin{pmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n & \mathbf{q}_{n+1} & \cdots & \mathbf{q}_m \end{pmatrix}$,

$$\mathbf{Q}'\mathbf{R}' = \begin{pmatrix} r_{11}\mathbf{q}_1 & r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 & \cdots & r_{1n}\mathbf{q}_1 + \cdots + r_{nn}\mathbf{q}_n \end{pmatrix} = \mathbf{A}.$$

Next for $\mathbf{Q}'$ to be orthogonal, the columns must form an orthonormal basis for $\mathbb{R}^m$. Hence, suffice to extend $T = \{\mathbf{q}_1, ..., \mathbf{q}_n\}$ to an orthonormal basis $\{\mathbf{q}_1, ..., \mathbf{q}_n, \mathbf{q}_{n+1}, ..., \mathbf{q}_m\}$ for $\mathbb{R}^m$.

Question 2(b) and (c)

In MATLAB, enter the following.

```
>> A=sym([1 1 0;1 1 0;1 1 1;0 1 1])
>> [Q R]=qr(A)
```

What is **Q** and **R** and explain how you might use the command `qr` in MATLAB to find a QR factorization of a $m \times n$ matrix **A**?

It is the full QR factorization of **A**. If we let

```
>> Q=Q(:,[1:3]), R=R([1:3],:)
```

Then we obtain the answer from tutorial 8 question 5.

In general, let $\mathbf{A} = \mathbf{Q}'\mathbf{R}'$ be the full QR factorization computed in MATLAB. Then let **Q** be the first n columns of $\mathbf{Q}'$, and **R** be the first n (nonzero) rows of $\mathbf{R}'$. The MATLAB code are

```
>> [Q R]=qr(A)
>> Q=(:,[1:n]), R=R([1:n],:)
```

Question 3(a)

Consider

$$p(\mathbf{X}) = \mathbf{X}^3 - 4\mathbf{X}^2 - \mathbf{X} + 4\mathbf{I}.$$

Compute $p(\mathbf{X})$ for $\mathbf{X} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}$.

```
>> X=[1 1 2;1 2 1;2 1 1]; X^3-4*X^2-X+4*eye(3)
```

Question 3(b)

Find the characteristic polynomial of **X**.

```
>> syms x; det(x*eye(3)-X)
```

Question 3(c)

Show that $\mathbf{X}$ invertible. Express the inverse of $\mathbf{X}$ as a function of $\mathbf{X}$.

$$\mathbf{X}^3 - 4\mathbf{X}^2 - \mathbf{X} + 4\mathbf{I} = \mathbf{0} \quad \Rightarrow \quad \mathbf{I} = -\frac{1}{4}(\mathbf{X}^3 - 4\mathbf{X}^2 - \mathbf{X}) = -\frac{1}{4}\mathbf{X}(\mathbf{X}^2 - 4\mathbf{X} - \mathbf{I})$$

This shows that $\mathbf{X}$ is invertible with inverse $-\frac{1}{4}(\mathbf{X}^2 - 4\mathbf{X} - \mathbf{I})$.

This question demonstrated the *Cayley-Hamilton theorem*, which states that if $p(x)$ is the characteristic polynomial of $\mathbf{X}$, then $p(\mathbf{X}) = 0$. This also show that if 0 is not an eigenvalue of $\mathbf{X}$, then the constant term of the characteristic polynomial $p(x)$ is nonzero, and we can use that to compute the inverse of $\mathbf{X}$.

Challenge: Show that if $\det(x\mathbf{I} - \mathbf{A}) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, then $a_0 = \det(\mathbf{A})$.

Hence, if $\det(A) \neq 0$, the Cayley-Hamilton theorem also shows that $\mathbf{A}$ is invertible with inverse $-\frac{1}{a_0}(\mathbf{A}^{n-1} + a_{n-1}\mathbf{A}^{n-2} + \cdots + a_1\mathbf{I})$.

Question 4(a)

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

Determine if $\mathbf{A}$ is diagonalizable. If $\mathbf{A}$ is diagonalizable, find an invertible $\mathbf{P}$ that diagonalizes $\mathbf{A}$ and determine $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

```
>> A=[1 -3 3;3 -5 3;6 -6 4]; syms x; simplify(det(x*eye(3)-A))
```

The eigenvalues are $\lambda = -2$ and $\lambda = 4$ with multiplicities $r_{-2} = 2, r_4 = 1$. Then $\mathbf{A}$ is diagonalizable if and only if the geometric multiplicity of eigenvalue $\lambda = -2$ is 2, $\dim(E_{-2}) = 2 = r_{-2}$.

`>> rref(-2*eye(3)-A)` It is indeed diagonalizable. Compute also the basis for eigenspace associated to $\lambda = 4$.

```
>> rref(4*eye(3)-A)
```

Let $\mathbf{P} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ and $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$.

Question 4(b)

$$\mathbf{A} = \begin{pmatrix} 9 & 8 & 6 & 3 \\ 0 & -1 & 3 & -4 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

```
>> A=[9 8 6 3;0 -1 3 -4;0 0 2 0;0 0 0 3]; rref(9*eye(4)-A)
>> rref(-1*eye(4)-A)
>> rref(2*eye(4)-A)
rref(3*eye(4)-A)
```

So, let $\mathbf{P} = \begin{pmatrix} 1 & -4 & -2 & 5 \\ 0 & 5 & 1 & -6 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}$ and $\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 9 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$

Question 4(c)

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

```
>> A=[1 0 0;1 1 0;0 1 1]; rref(eye(3)-A)
```

Question 4(d)

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

```
>> A=[0 0 1 0;0 0 0 1;1 0 0 0;0 1 0 0]; simplify(det(x*eye(4)-A))
>> rref(eye(4)-A)
>> rref(-eye(4)-A)
```

Hence, $\mathbf{A}$ is diagonalizable, with $\mathbf{P} = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Question 4(e)

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 1 & -1 \\ -4 & 2 & 3 \end{pmatrix}.$$

```
>> A=[-1 1 1;1 1 -1;-4 2 3]; simplify(det(x*eye(3)-A))
>> solve(ans)
```

Question 5(a) and (b)

(a) Show that λ is an eigenvalue of $\mathbf{A}$ if and only if it is an eigenvalue of $\mathbf{A}^T$.

$$\det(x\mathbf{I} - \mathbf{A}) = \det((x\mathbf{I} - \mathbf{A})^T) = \det((x\mathbf{I})^T - \mathbf{A}^T) = \det(x\mathbf{I} - \mathbf{A}^T).$$

Hence the roots of $\det(x\mathbf{I} - \mathbf{A})$ are exactly the roots of $\det(x\mathbf{I} - \mathbf{A}^T)$.

(b) Suppose $\mathbf{A}$ is diagonalizable. Is $\mathbf{A}^T$ diagonalizable? Justify your answer.

Yes. Write $\mathbf{A} = \mathbf{PDP}^{-1}$. Then

$$\mathbf{A}^T = (\mathbf{PDP}^{-1})^T = (\mathbf{P}^{-1})^T\mathbf{D}^T\mathbf{P}^T = \mathbf{QDQ}^{-1},$$

where the last equality follows from the fact that $\mathbf{D}$ is diagonal and thus symmetric, and letting $\mathbf{Q} = (\mathbf{P}^{-1})^T$ (recall that $(\mathbf{P}^{-1})^T = (\mathbf{P}^T)^{-1}$). That is, $\mathbf{Q} = (\mathbf{P}^{-1})^T$ diagonalizes $\mathbf{A}^T$.

Question 5(c) and (d)

(c) Suppose $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ associated to eigenvalue λ. Show that $\mathbf{v}$ is an eigenvector of $\mathbf{A}^k$ associated to eigenvalue λ^k for any positive integer k.

By definition, we have $\mathbf{Av} = \lambda\mathbf{v}$. Then

$$\mathbf{A}^k\mathbf{v} = \mathbf{A}^{k-1}\mathbf{Av} = \lambda\mathbf{A}^{k-2}\mathbf{Av} = \lambda^2\mathbf{A}^{k-3}\mathbf{Av} = \cdots = \lambda^{k-1}\mathbf{Av} = \lambda^k\mathbf{v}.$$

Since $\mathbf{v} \neq \mathbf{0}$, $\mathbf{v}$ is a witness to λ^k being an eigenvalue of $\mathbf{A}^k$.

(d) If $\mathbf{A}$ is invertible, show that $\mathbf{v}$ is an eigenvector of $\mathbf{A}^k$ associated to eigenvalue λ^k for any negative integer k.

Suppose $k = -1$. First note that since $\mathbf{A}$ is invertible, $k \neq 0$. Then $\mathbf{Av} = \lambda\mathbf{v} \iff \lambda^{-1}\mathbf{v} = \mathbf{A}^{-1}\mathbf{v}$. Hence, λ^{-1} is an eigenvalue of $\mathbf{A}^{-1}$. The rest of the argument follows from (c).

Question 5(e), (f), and (g)

(e) A square matrix is said to be *nilpotent* if there is a positive integer k such that $\mathbf{A}^k = \mathbf{0}$. Show that if $\mathbf{A}$ is nilpotent, then 0 is the only eigenvalue.

Let λ be an eigenvalue of $\mathbf{A}$ and $\mathbf{v}$ be an eigenvector associated to λ. By (c), $\mathbf{0} = \mathbf{A}^k\mathbf{v} = \lambda^k\mathbf{v}$. Since $\mathbf{v} \neq \mathbf{0}$, necessarily $\lambda^k = 0$, and hence $\lambda = 0$.

(f) Let $\mathbf{A}$ be a $n \times n$ matrix with one eigenvalue λ with algebraic multiplicity n. Show that $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ is a scalar matrix, $\mathbf{A} = \lambda\mathbf{I}$.

Suppose $\mathbf{A}$ is diagonalizable, say $\mathbf{A} = \mathbf{PDP}^{-1}$ for some invertible matrix $\mathbf{P}$. Now since λ is the only eigenvalue with multiplicity n, necessarily $\mathbf{D} = \lambda\mathbf{I}$. Hence,

$$\mathbf{A} = \mathbf{P}\lambda\mathbf{I}\mathbf{P}^{-1} = \lambda\mathbf{P}\mathbf{I}\mathbf{P}^{-1} = \lambda\mathbf{I},$$

that is, $\mathbf{A} = \lambda\mathbf{I}$ is a scalar matrix. It is clear that a scalar matrix is diagonalizable.

(g) Show that the only diagonalizable nilpotent matrix is the zero matrix.

Let $\mathbf{A}$ be a nilpontent matrix. Then 0 is the only eigenvalue. If $\mathbf{A}$ is diagonalizable, then $\mathbf{A} = \mathbf{P}\,\mathrm{diag}(0, 0, ..., 0)\mathbf{P}^{-1} = \mathbf{0}$ for some invertible matrix P.