Advanced Linear Algebra

Hiroshi SUZUKI*
Division of Natural Sciences
International Christian University

June 14, 2008

*E-mail:hsuzuki@icu.ac.jp

Solutions to Homework 1

6.1.21 *Show that the following identity holds for vectors in any inner product space.*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \frac{1}{4}\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \frac{1}{4}\|\boldsymbol{u}-\boldsymbol{v}\|^2.$$

Solution.

$$\begin{aligned}
&\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2 \\
&= \langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v} \rangle - \langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{u}-\boldsymbol{v} \rangle \\
&= (\langle \boldsymbol{u}, \boldsymbol{u} \rangle + \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \langle \boldsymbol{v}, \boldsymbol{u} \rangle + \langle \boldsymbol{v}, \boldsymbol{v} \rangle) - (\langle \boldsymbol{u}, \boldsymbol{u} \rangle - \langle \boldsymbol{u}, \boldsymbol{v} \rangle - \langle \boldsymbol{v}, \boldsymbol{u} \rangle + \langle \boldsymbol{v}, \boldsymbol{v} \rangle) \\
&= 4\langle \boldsymbol{u}, \boldsymbol{v} \rangle.
\end{aligned}$$

Hence we have the formula. ■

6.1.23 *Let $\boldsymbol{p} = p(x)$ and $\boldsymbol{q} = q(x)$ be polynomials in P_2. Show that*

$$\langle \boldsymbol{p}, \boldsymbol{q} \rangle = p(0)q(0) + p(\frac{1}{2})q(\frac{1}{2}) + p(1)q(1)$$

is an inner product on P_2. Is this an inner product on P_3? Explain.

Solution. Clearly for all $\boldsymbol{p}, \boldsymbol{q} \in P_2$ $\langle \boldsymbol{p}, \boldsymbol{q} \rangle$ is a real number.

1. $\langle \boldsymbol{p}, \boldsymbol{q} \rangle = p(0)q(0) + p(\frac{1}{2})q(\frac{1}{2}) + p(1)q(1) = q(0)p(0) + q(\frac{1}{2})p(\frac{1}{2}) + q(1)p(1) = \langle \boldsymbol{q}, \boldsymbol{p} \rangle$ and the property 1 holds.

2. Let $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{q} \in P_2$. Then

$$\begin{aligned}
\langle \boldsymbol{p}_1 + \boldsymbol{p}_2, \boldsymbol{q} \rangle &= (p_1(0) + p_2(0)q(0) + (p_1(\frac{1}{2}) + p_2(\frac{1}{2}))q(\frac{1}{2}) + (p_(1) + p_2(1))q(1) \\
&= p_1(0)q(0) + p_1(\frac{1}{2})q(\frac{1}{2}) + p_1(1)q(1) + p_2(0)q(0) + p_2(\frac{1}{2})q(\frac{1}{2}) + p_2(1)q(1) \\
&= \langle \boldsymbol{p}_1, \boldsymbol{q} \rangle + \langle \boldsymbol{p}_2, \boldsymbol{q} \rangle.
\end{aligned}$$

   Thus the property 2 holds.

3. For all real k,

$$\langle k\boldsymbol{p}, \boldsymbol{q} \rangle = kp(0)q(0) + kp(\frac{1}{2})q(\frac{1}{2}) + kp(1)q(1) = k(p(0)q(0) + p(\frac{1}{2})q(\frac{1}{2}) + p(1)q(1)) = k\langle \boldsymbol{p}, \boldsymbol{q} \rangle.$$

   Hence the property 3 holds.

4. $\langle \boldsymbol{p}, \boldsymbol{p} \rangle = p(0)p(0) + p(\frac{1}{2})p(\frac{1}{2}) + p(1)p(1) \geq 0$. Suppose $\langle \boldsymbol{p}, \boldsymbol{p} \rangle = 0$. Then $p(0) = p(\frac{1}{2}) = p(1) = 0$. Since $\boldsymbol{p} = p(x)$ is of degree at most 2, this implies $\boldsymbol{p} = p(x) = 0$. Clearly if $\boldsymbol{p} = \boldsymbol{0}$, $\langle \boldsymbol{p}, \boldsymbol{p} \rangle = 0$. The last part is also a consequence of Theorem 6.1.1 (a).

Finally if we allow $\boldsymbol{p} = p(x) \in P_3$, then $\boldsymbol{0} \neq \boldsymbol{p} = p(x) = x(x - \frac{1}{2})(x-1) \in P_3$ but $\langle \boldsymbol{p}, \boldsymbol{p} \rangle = 0$. Hence $\langle \boldsymbol{p}, \boldsymbol{q} \rangle$ cannot be an inner product on P_3 as it does not satisfy the property 4. ■

6.2.21 *Let V be an inner product space. Show that if $\boldsymbol{w}$ is orthogonal to each of the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r$, then it if orthogonal to every vector in* $\operatorname{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$.

Solution. Let $\boldsymbol{v} \in \operatorname{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$. Then there exist real numbers $c_1, c_2, \ldots, c_r$ such that $\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_r\boldsymbol{u}_r$. Hence

$$\langle \boldsymbol{w}, \boldsymbol{v} \rangle = \langle \boldsymbol{w}, c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_r\boldsymbol{u}_r \rangle = c_1\langle \boldsymbol{w}, \boldsymbol{u}_1 \rangle + c_2\langle \boldsymbol{w}, \boldsymbol{u}_2 \rangle + \cdots + c_r\langle \boldsymbol{w}, \boldsymbol{u}_r \rangle.$$

By assumption, $\langle \boldsymbol{w}, \boldsymbol{v} \rangle = 0$, and we have the assertion. ■

Solutions to Homework 2

Main Results (6.3.4, 6.3.5, 6.3.6, 6.4.4) *Let V be an n-dimensional inner product space, and W a subspace of W. Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ a basis of W, $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m\}$ an orthonormal basis of W and $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m, \boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n\}$ be an orthonormal basis of V. Let $A = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m]$. Then the following hold.*

(i) $W^\perp = \{\boldsymbol{v} \mid \langle \boldsymbol{v}, \boldsymbol{w} \rangle = 0$ *for all* $\boldsymbol{w} \in W\} = \mathrm{Span}(\boldsymbol{q}_{m+1}, \ldots, \boldsymbol{q}_n)$.

(ii) *For $\boldsymbol{v} \in V$, there exist $\boldsymbol{w} \in W$ and $\boldsymbol{w}' \in W^\perp$ such that $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{w}'$. Moreover, $\boldsymbol{w} \in W$ and $\boldsymbol{w}' \in W^\perp$ are uniquely determined and $\boldsymbol{w} = \mathrm{proj}_W(\boldsymbol{v})$.*

(iii) $\mathrm{proj}_W(\boldsymbol{v}) = \langle \boldsymbol{v}, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 + \langle \boldsymbol{v}, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2 + \cdots + \langle \boldsymbol{v}, \boldsymbol{q}_m \rangle \boldsymbol{q}_m$ *for all* $\boldsymbol{v} \in V$.

(iv) *Let $1 \le i \le m-1$. Suppose $W_i = \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i) = \mathrm{Span}(\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_i)$. Let*

$$\boldsymbol{p} = \boldsymbol{u}_{i+1} - \mathrm{proj}_{W_i}(\boldsymbol{u}_{i+1}) = \boldsymbol{u}_{i+1} - \langle \boldsymbol{u}_{i+1}, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 - \cdots - \langle \boldsymbol{u}_{i+1}, \boldsymbol{q}_i \rangle \boldsymbol{q}_i, \ \textit{and}\ \boldsymbol{q} = \frac{1}{\|\boldsymbol{p}\|}\boldsymbol{p}.$$

Then $W_{i+1} = \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i, \boldsymbol{u}_{i+1}) = \mathrm{Span}(\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_i, \boldsymbol{q})$ and $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_i, \boldsymbol{q}\}$ is an orthonormal basis of W_{i+1} and $\boldsymbol{q}$ can be chosen for $\boldsymbol{q}_{i+1}$.

Moreover, if V is the Euclidean inner product space on $\boldsymbol{R}^m$, then

(v) $\mathrm{proj}_W(\boldsymbol{v}) = A(A^TA)^{-1}A^T\boldsymbol{v}$ *for all $\boldsymbol{v} \in V$. In particular, if $\boldsymbol{u}_1 = \boldsymbol{q}_1$, $\boldsymbol{u}_2 = \boldsymbol{q}_2, \ldots, \boldsymbol{u}_m = \boldsymbol{q}_m$, then $A^TA = I_m$ and $\mathrm{proj}_W(\boldsymbol{v}) = AA^T\boldsymbol{v}$, which is equal to the expression in* (iii).

6.3.18 *Let $\boldsymbol{R}^4$ have the Euclidean inner product. Use the Gram-Schmidt process to transform the basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ into an orthonormal basis.*

$$\boldsymbol{u}_1 = (0,2,1,0),\ \boldsymbol{u}_2 = (1,-1,0,0),\ \boldsymbol{u}_3 = (1,2,0,-1),\ \boldsymbol{u}_4 = (1,0,0,1).$$

Solution. Our notation is a bit different from the one in Example 7 (page 325). Let $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$, $\boldsymbol{q}_4$ are the orthonormal basis obtained by the Gram-Schmidt process.

Step 1. First let $\boldsymbol{v}_1 = \boldsymbol{u}_1$. Then

$$\boldsymbol{q}_1 = \frac{1}{\|\boldsymbol{v}_1\|}\boldsymbol{v}_1 = \frac{1}{\sqrt{5}}(0,2,1,0).$$

Step 2. Let $W_1 = \mathrm{Span}(\boldsymbol{q}_1)$, and $\boldsymbol{v}_2 = \boldsymbol{u}_2 - \mathrm{proj}_{W_1}(\boldsymbol{u}_2)$

$$\begin{aligned}
\boldsymbol{v}_2 &= \boldsymbol{u}_2 - \langle \boldsymbol{u}_2, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 \\
&= (1,-1,0,0) - \langle (1,-1,0,0), \frac{1}{\sqrt{5}}(0,2,1,0) \rangle \frac{1}{\sqrt{5}}(0,2,1,0) \\
&= \frac{1}{5}(5(1,-1,0,0) + 2(0,2,1,0)) = \frac{1}{5}(5,-1,2,0). \\
\boldsymbol{q}_2 &= \frac{1}{\|(5,-1,2,0)\|}(5,-1,2,0) = \frac{1}{\sqrt{30}}(5,-1,2,0).
\end{aligned}$$

Spep 3. Let $W_2 = \mathrm{Span}(\boldsymbol{q}_1, \boldsymbol{q}_2)$, and $\boldsymbol{v}_3 = \boldsymbol{u}_3 - \mathrm{proj}_{W_2}(\boldsymbol{u}_3)$

$$\begin{aligned}
\boldsymbol{v}_3 &= \boldsymbol{u}_3 - \langle \boldsymbol{u}_3, \boldsymbol{q}_1 \rangle \boldsymbol{q}_1 - \langle \boldsymbol{u}_3, \boldsymbol{q}_2 \rangle \boldsymbol{q}_2 \\
&= (1,2,0,-1) - \langle (1,2,0,-1), \frac{1}{\sqrt{5}}(0,2,1,0) \rangle \frac{1}{\sqrt{5}}(0,2,1,0) \\
&\quad - \langle (1,2,0,-1), \frac{1}{\sqrt{30}}(5,-1,2,0) \rangle \frac{1}{\sqrt{30}}(5,-1,2,0) \\
&= \frac{1}{30}(30(1,2,0,-1) - 6\cdot 4(0,2,1,0) - 3(5,-1,2,0)) = \frac{1}{30}(15,15,-30,-30). \\
\boldsymbol{q}_3 &= \frac{1}{\|(1,1,-2,-2)\|}(1,1,-2,-2) = \frac{1}{\sqrt{10}}(1,1,-2,-2).
\end{aligned}$$

Spep 4. Let $W_3 = \text{Span}(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3)$, and $\boldsymbol{v}_4 = \boldsymbol{u}_4 - \text{proj}_{W_3}(\boldsymbol{u}_4)$

$$\begin{aligned}
\boldsymbol{v}_4 &= \boldsymbol{u}_4 - \langle \boldsymbol{u}_4, \boldsymbol{q}_1\rangle \boldsymbol{q}_1 - \langle \boldsymbol{u}_4, \boldsymbol{q}_2\rangle \boldsymbol{q}_2 - \langle \boldsymbol{u}_4, \boldsymbol{q}_3\rangle \boldsymbol{q}_3 \\
&= (1,0,0,1) - \langle (1,0,0,1), \frac{1}{\sqrt{5}}(0,2,1,0)\rangle \frac{1}{\sqrt{5}}(0,2,1,0) \\
&\quad - \langle (1,0,0,1), \frac{1}{\sqrt{30}}(5,-1,2,0)\rangle \frac{1}{\sqrt{30}}(5,-1,2,0) \\
&\quad - \langle (1,0,0,1), \frac{1}{\sqrt{10}}(1,1,-2,-2)\rangle \frac{1}{\sqrt{10}}(1,1,-2,-2) \\
&= \frac{1}{30}(30(1,0,0,1) - 0 - 5(5,-1,2,0) - 3\cdot(-1)(1,1,-2,-2)) = \frac{1}{30}(8,8,-15,24). \\
\boldsymbol{q}_4 &= \frac{1}{\|(1,1,-2,3)\|}(1,1,-2,3) = \frac{1}{\sqrt{15}}(1,1,-2,3).
\end{aligned}$$

Therefore,

$$\boldsymbol{q}_3 = \frac{1}{\sqrt{5}}(0,2,1,0),\ \boldsymbol{q}_2 = \frac{1}{\sqrt{30}}(5,-1,2,0),\ \boldsymbol{q}_3 = \frac{1}{\sqrt{10}}(1,1,-2,-2),\ \boldsymbol{q}_4 = \frac{1}{\sqrt{15}}(1,1,-2,3).$$

■

6.4.4 *Find the orthogonal projection of $\boldsymbol{u}$ onto the subspace of $\boldsymbol{R}^3$ spanned by the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$.*

(a) $\boldsymbol{u} = (2,1,3)$; $\boldsymbol{v}_1 = (1,1,0)$, $\boldsymbol{v}_2 = (1,2,1)$.

(b) $\boldsymbol{u} = (1,-6,1)$; $\boldsymbol{v}_1 = (-1,2,1)$, $\boldsymbol{v}_2 = (2,2,4)$.

Solution. Let

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} -1 & 2 \\ 2 & 2 \\ 1 & 4 \end{bmatrix}. \text{ Then } A^TA = \begin{bmatrix} 2 & 3 \\ 3 & 6 \end{bmatrix},\ B^TB = \begin{bmatrix} 6 & 6 \\ 6 & 24 \end{bmatrix}.$$

Hence

$$A(A^TA)^{-1}A^T = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 6 & -3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}, \text{ and}$$

$$B(B^TB)^{-1}B^T = \frac{1}{18}\begin{bmatrix} -1 & 2 \\ 2 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 4 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 & 1 \\ 2 & 2 & 4 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 4 & -2 & 2 \\ -2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}.$$

Therefore

$$\frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ 7 \\ 5 \end{bmatrix}, \text{ and } \frac{1}{6}\begin{bmatrix} 4 & -2 & 2 \\ -2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ -6 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \\ -1 \end{bmatrix}.$$

(a) Let $W = \text{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$. Then by Theorem 6.4.4, $\text{proj}(\boldsymbol{u}) = \frac{1}{3}(2,7,5)$.

(b) Let $W = \text{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$. Then by Theorem 6.4.4, $\text{proj}(\boldsymbol{u}) = (3,-4,-1)$

■

6.4.10 *Let W be the plane with equation $5x - 3y + z = 0$.*

(a) *Find a basis for W.*

(b) *Use Formula* (6) *to find the standard matrix for the orthogonal projection onto W.*

(c) *Use the matrix obtained in* (b) *to find the orthogonal projection of a point $P_0(x_0, y_0, z_0)$ onto W.*

(d) *Find the distance between the point $P_0(1, -2, 4)$ and the plane W, and check your result using* Theorem 3.5.2.

Solution.

(a) W is the solution set of a system of linear equation with its augmented matrix $[5, -3, 1, 0]$. Hence

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = s \begin{bmatrix} \frac{3}{5} \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{1}{5} \\ 0 \\ 1 \end{bmatrix}, \text{ and a set of basis of } W \text{ is } \left\{ \begin{bmatrix} 3 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix} \right\}.$$

Here s and t are parameters. (Note that all elements of W can be written as a linear combination of these two vectors. It is clear by looking at the second and the third entries that these two vectors are linearly independent.) ■

(b) Let A be a matrix consisting of basis vectors of W obtained in (a).

$$[P] = A(A^TA)^{-1}A^T = \frac{1}{811} \begin{bmatrix} 3 & -1 \\ 5 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 25 & 3 \\ 3 & 25 \end{bmatrix} \begin{bmatrix} 3 & 5 & 0 \\ -1 & 0 & 5 \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 10 & 15 & -5 \\ 15 & 26 & 3 \\ -5 & 3 & 34 \end{bmatrix}.$$

(c) Let $\boldsymbol{x} = (x_0, y_0, z_0)^T$. Then

$$[P]\boldsymbol{x} = \frac{1}{35} \begin{bmatrix} 10x_0 + 15y_0 - 5z_0 \\ 15x_0 + 26y_0 + 3z_0 \\ -5x_0 + 3y_0 + 34z_0 \end{bmatrix}$$

(d) $[P](1, -2, 4)^T = \frac{1}{7}(-8, -5, 25)$. By Theorem 3.5.2

$$D = \frac{|5x_0 - 3y_0 + z_0 + 0|}{\sqrt{25 + 9 + 1}} = \frac{5 + 6 + 4}{\sqrt{35}} = \frac{3\sqrt{35}}{7}.$$

On the other hand

$$\|(1, -2, 4) - \text{proj}(1, -2, 4)\| = \|(1, -2, 4) - \frac{1}{7}(-8, -5, 25)\| = \frac{1}{7}\|(15, -9, 3)\| = \frac{3\sqrt{35}}{7}.$$

Hence these two values agree. ■

Solutions to Homework 3

6.5.9 *Consider the basis $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ and $B' = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ for $\boldsymbol{R}^3$, where*

$$\boldsymbol{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \boldsymbol{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} -5 \\ 8 \\ -5 \end{bmatrix}$$

(a) *Find the transition matrix from B to B'.*

(b) *Compute the coordinate vector $[\boldsymbol{w}]_B$ and use (9) to compute $[\boldsymbol{w}]_{B'}$.*

(c) *Check your work by computing $[\boldsymbol{w}]_{B'}$ directly.*

Solution.

(a) Let P be the transition matrix from B to B'. Then $P = [[\boldsymbol{u}_1]_{B'}, [\boldsymbol{u}_2]_{B'}, [\boldsymbol{u}_3]_{B'}]$. Hence it suffices to express $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ as a linear combination of $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$.

$$\begin{aligned}
\boldsymbol{u}_1 &= \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 3\begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + 5\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} = 3\boldsymbol{v}_1 - 2\boldsymbol{v}_2 + 5\boldsymbol{v}_3, \text{ and } [\boldsymbol{u}_1]_{B'} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix}, \\
\boldsymbol{u}_2 &= \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} = 2\boldsymbol{v}_1 - 3\boldsymbol{v}_2 + \boldsymbol{v}_3, \text{ and } [\boldsymbol{u}_2]_{B'} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \\
\boldsymbol{u}_3 &= \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \frac{5}{2}\begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + 6\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} = \frac{5}{2}\boldsymbol{v}_1 - \frac{1}{2}\boldsymbol{v}_2 + 6\boldsymbol{v}_3, \text{ and } [\boldsymbol{u}_3]_{B'} = \begin{bmatrix} 5/2 \\ -1/2 \\ 6 \end{bmatrix}.
\end{aligned}$$

Therefore,

$$P = [[\boldsymbol{u}_1]_{B'}, [\boldsymbol{u}_2]_{B'}, [\boldsymbol{u}_2]_{B'}] = \begin{bmatrix} 3 & 2 & 5/2 \\ -2 & -3 & -1/2 \\ 5 & 1 & 6 \end{bmatrix}.$$

(b) First express $\boldsymbol{w}$ as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$.

$$\boldsymbol{w} = \begin{bmatrix} -5 \\ 8 \\ -5 \end{bmatrix} = 9\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} - 9\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} - 5\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 9\boldsymbol{u}_1 - 9\boldsymbol{u}_2 - 5\boldsymbol{u}_3. \text{ and } [\boldsymbol{w}]_B = \begin{bmatrix} 9 \\ -9 \\ -5 \end{bmatrix}.$$

Since P is the transition matrix from B to B',

$$[\boldsymbol{w}]_{B'} = P[\boldsymbol{w}]_B = \begin{bmatrix} 3 & 2 & 5/2 \\ -2 & -3 & -1/2 \\ 5 & 1 & 6 \end{bmatrix}\begin{bmatrix} 9 \\ -9 \\ -5 \end{bmatrix} = \begin{bmatrix} -7/2 \\ 23/2 \\ 6 \end{bmatrix}.$$

(c) Now we find $[\boldsymbol{w}]_{B'}$.

$$\boldsymbol{w} = \begin{bmatrix} -5 \\ 8 \\ -5 \end{bmatrix} = -\frac{7}{2}\begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix} + \frac{23}{2}\begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + 6\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} = -\frac{7}{2}\boldsymbol{v}_1 + \frac{23}{2}\boldsymbol{v}_2 + 6\boldsymbol{v}_3.$$

■

基底変換解説 6.5.9 を例にとります。基底 $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ に対して $P_B = [\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]$ を 3 次正方行列とします。同じように $P_{B'} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3]$ と定義します。$\boldsymbol{R}^3$ のベクトル $\boldsymbol{w}$ を B の一次結合に表すと、$\boldsymbol{w} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + x_3\boldsymbol{u}_3$ だとすると、これは、行列で

$$\boldsymbol{w} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + x_3\boldsymbol{u}_3 = [\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = P_B[\boldsymbol{w}]_B$$

とも表すことができます。標準基底を $B_0 = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ とすると、$\boldsymbol{w} = [\boldsymbol{w}]_{B_0}$ ですから、$[\boldsymbol{w}]_{B_0} = P_B[\boldsymbol{w}]_B$ と書くこともできます。同様にして、$[\boldsymbol{w}]_{B_0} = P_{B'}[\boldsymbol{w}]_{B'}$ と書くことができます。さて、B は基底で、$\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ は一次独立でしたから、P_B は可逆行列です。したがって、$[\boldsymbol{w}]_B = P_B^{-1}\boldsymbol{w}$, $[\boldsymbol{w}]_{B'} = P_{B'}^{-1}\boldsymbol{w}$ とも書くことができます。では、B から B' への基底変換行列 (the transition matrix from B to B') $P = P_{B,B'}$ はどうなるでしょうか。定義から

$$P = [[\boldsymbol{u}_1]_{B'}, [\boldsymbol{u}_2]_{B'}, [\boldsymbol{u}_3]_{B'}] = [P_{B'}^{-1}\boldsymbol{u}_1, P_{B'}^{-1}\boldsymbol{u}_2, P_{B'}^{-1}\boldsymbol{u}_3] = P_{B'}^{-1}[\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3] = P_{B'}^{-1}P_B$$

となります。すなわち、P_B^{-1} を求めれば、それと、$P_{B'}$ とかけて、P を求めることができます。また、

$$[\boldsymbol{w}]_{B'} = P_{B'}^{-1}\boldsymbol{w} = P_{B'}^{-1}P_B[\boldsymbol{w}]_B = P[\boldsymbol{w}]_B$$

も分かります。

では、直接的に P を求めるときは、どうするのでしょうか。これは、$[\boldsymbol{u}_1]_{B'}$, $[\boldsymbol{u}_2]_{B'}$, $[\boldsymbol{u}_3]_{B'}$ を求めるのですから、$P_{B'}\boldsymbol{x} = \boldsymbol{u}_1$, $P_{B'}\boldsymbol{x} = \boldsymbol{u}_2$, $P_{B'}\boldsymbol{x} = \boldsymbol{v}_3$ と三つの方程式を解いて、求めるわけです。拡大係数行列 $[P_{B'}, \boldsymbol{u}_1]$, $[P_{B'}, \boldsymbol{u}_2]$, $[P_{B'}, \boldsymbol{u}_3]$ からスタートして、$P_{B'}$ の部分が単位行列になったときに、$[\boldsymbol{u}_1]_{B'}$, $[\boldsymbol{u}_2]_{B'}$, $[\boldsymbol{u}_3]_{B'}$ が求まることになります。しかしそうであれば、三ついっぺんに、$[P_{B'}, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]$ という 3×6 行列を考え、ここから、既約ガウス行列にもっていけば、$[I, [\boldsymbol{u}_1]_{B'}, [\boldsymbol{u}_2]_{B'}, [\boldsymbol{u}_3]_{B'}]$ となるわけです。これは、何かを思い出しますね。逆行列の求め方です。ただ、逆行列の時には、最初、$[A, I]$ というような形からスタートするのでした。しかし、我々のスタートは、$[P_{B'}, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3]$ でした。しかし、この右半分は、P_B ですから、$[P_{B'}, P_B]$ がスタート地点だとも言えます。ここで、P_B^{-1} をこの行列にかけてみましょう。すると、$P_B^{-1}[P_{B'}, P_B] = [P_B^{-1}P_{B'}, I]$ となります。ちょっと、長くなりましたが、これを、変形していくと、その既約ガウス行列が、$[I, [\boldsymbol{u}_1]_{B'}, [\boldsymbol{u}_2]_{B'}, [\boldsymbol{u}_3]_{B'}]$ になったとも言うことができるわけです。ということは、$[\boldsymbol{u}_1]_{B'}, [\boldsymbol{u}_2]_{B'}, [\boldsymbol{u}_3]_{B'}] = (P_B^{-1}P_{B'})^{-1} = P_{B'}^{-1}P_B$ となり、上で求めた、$P = P_{B'}^{-1}P_B$ と同じ式が出ます。すなわち、本質的には同じことをしていることになります。あとは、どちらが、計算が簡単か、間違えにくいかということです。

私はどうするかというと、このような記号を使って、すべての問題の答えを表してみます。そのあとで、どれを計算するのが、合理的かを判断します。すなわち、答えは、(a) $P = P_{B'}^{-1}P_B$, (b) $[\boldsymbol{w}]_B = P_B^{-1}\boldsymbol{w}$, $[\boldsymbol{w}]_{B'} = P[\boldsymbol{w}]_B$. (c) $[\boldsymbol{w}]_{B'} = P_{B'}^{-1}\boldsymbol{w}$ となります。

どうでしょうか。このように分かったときに、何を計算しますか。私なら、$P_{B'}^{-1}$ を計算します。

最初の $\boldsymbol{w} = P_B[\boldsymbol{w}]_B$ に戻ってみましょう。これは、なにか変な感じもしますが、行列 P_B をブロック（この場合は列）にわけて、行列の積を計算するという手法ですね。何もへんな事をしているわけではありません。なかなか有効な方法ですので、慣れておくと良いですよ。

6.6.10 *Let a rectangular $x'y'z'$-coordinate system be obtained by rotating a rectangular xyz-coordinate system counterclockwise about the x-axis (looking along the positive x-axis toward the origin) through the angle $\theta = 3\pi/4$.*

(a) *Find the $x'y'z'$-coordinates of the point whose xyz-coordinates are* $(-1, 2, 5)$.

(b) *Find the x, y, z-coordinates of the point whose $x'y'z'$-coordinates are* $(1, 6, -3)$.

Solution. Let T be the rotation around x-axis and B the standard basis. Then the general form of the transition matrix is

$$[T] = [[T(\boldsymbol{e}_1)]_B, [T(\boldsymbol{e}_2)]_B, [T(\boldsymbol{e}_3)]_B] = P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix},\ P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}$$

(a) Since $\theta = 3\pi/4$, $\cos\theta = -\sqrt{2}/2$ and $\sin\theta = \sqrt{2}/2$. Hence

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \\ 0 & -\sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} -1 \\ 3\sqrt{2}/2 \\ -7\sqrt{2}/2 \end{bmatrix}.$$

(b) Now we use P.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ -3\sqrt{2}/2 \\ 9\sqrt{2}/2 \end{bmatrix}.$$

座標軸の回転　注意しなければならないのは、これは、点を回転させるのではなく、座標軸を回転させ、回転させた座標軸のもとで、回転する前の点を表すというものです。そこで、基底変換行列の逆行列を用いることになることに注意して下さい。Example 4, Example 5 を参照して下さい。これから考えたいのは、このような x-軸や、y-軸、z-軸に関する回転ではなく、一般の方向ベクトルを軸とした回転は、どう表されるのか、逆に、直交変換は、回転 (Rotation) や 鏡映 (Reflection) その合成などで表せるのかといったことです。直交変換をある基底を用いると、対角行列にできたり、対称行列も、ある基底で対角行列にできたりといったことを考えます。それには、別の基底のもとで、一次変換がどう表されるかを考える、それが内積空間のときは、どうだろうかなどなど、が次の目標です。

6.6.15　*Exercise 6.6.14 states that a* 2×2 *orthogonal matrix* A *has one of two possible forms:*

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ or } A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

where $0 \leq \theta < 2\pi$.

(a) *Use the result in Exercise 14 to prove that multiplication by a* 2×2 *orthogonal matrix is either a rotation or a rotation followed by a reflection about the* x*-axis.*

(b) *Show that multiplication by* A *is a ration if* $\det(A) = 1$ *and that a rotation followed by a reflection if* $\det(A) = -1$.

Solution.

(a) Clearly the first matrix is a rotation through an angle θ. On the other hand

$$\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}.$$

The last matrix is clearly a rotation through an angle $-\theta$.

(b) This is clear because two matrices above are characterized by the value of $\det(A)$, and the $\det(A) = 1$ for the first matrix and $\det(A) = -1$ for the second. ■

2 次の直交行列　Exercise 6.6.14 が重要ですね。宿題には、ちょっと大変だと思い出しませんでした。この問題の様に、直交行列は、回転か、回転のあと x-軸についての鏡映（対称変換）との合成になっています。実は、このあとの方は、ある軸に関する鏡映になっています。その軸を見つけることができますか。また、どんな、回転も二つの鏡映の積に書くことができます。それも、幾何的に考えると簡単ですから、考えてみて下さい。最後に英語を一つ。a rotation followed by a reflection は、まず回転をしてそれから、鏡映ですね。しかし、事実としては、回転をしてから鏡映をしたものは、鏡映をしてから回転をしたものと同じになります。ただし、回転の角度は変わります。

計算を確かめるためにもコンピュータやネットを使うのは良いと思いますよ。手計算も大切ですが、確かめることも大切です。私の利用するサイトをあげておきます。http://www.math.odu.edu/ bogacki/cgi-bin/lat.cgi, http://wims.unice.fr/wims/, http://www.quickmath.com/ これらでは、Gram-Schmidt は直接的にはできませんが、それが計算できるサイトもありますね。本格的な代数計算ができるのは、magma (http://magma.maths.usyd.edu.au/calc/) ですが、最初はちょっと難しいかな。

6.6.9 *Let a rectangular $x'y'z'$-coordinate system be obtained by rotating a rectangular xyz-coordinate system counterclockwise about the y-axis (looking along the positive y-axis toward the origin) through the angle $\theta = \pi/3$.*

(a) *Find the $x'y'z'$-coordinates of the point whose xyz-coordinates are $(-1, 2, 5)$.*

(b) *Find the x, y, z-coordinates of the point whose $x'y'z'$-coordinates are $(1, 6, -3)$.*

Solution. Let T be the rotation around y-axis and B the standard basis. Then the general form of the transition matrix is

$$[T] = [[T(\boldsymbol{e}_1)]_B, [T(\boldsymbol{e}_2)]_B, [T(\boldsymbol{e}_3)]_B] = P = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}, \; P^{-1} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

(a) Since $\theta = \pi/3$, $\cos\theta = 1/2$ and $\sin\theta = \sqrt{3}/2$. Hence

$$\begin{bmatrix} 1/2 & 0 & \sqrt{3}/2 \\ 0 & 1 & 0 \\ -\sqrt{3}/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} (-1+5\sqrt{3})/2 \\ 2 \\ (5+\sqrt{3})/2 \end{bmatrix}.$$

(b) Now we use P.

$$\begin{bmatrix} 1/2 & 0 & -\sqrt{3}/2 \\ 0 & 1 & 0 \\ \sqrt{3}/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \\ -3 \end{bmatrix} = \begin{bmatrix} (1+3\sqrt{3})/2 \\ 6 \\ (-3+\sqrt{3})/2 \end{bmatrix}.$$

■

Diagonalization

Definition 1 A square matrix A is called *diagonalizable* if there is an invertible matrix P such that $P^{-1}AP$ is a diagonal matirx; the matrix P is said to *diangonalize* A.

Theorem 1 (7.2.1) *If A is an $n \times n$ matrix, then the following are equivalent.*

(a) *A is diagonalizable.*

(b) *A has n linearly independent eigenvectors.*

Proof. (a)$\Rightarrow$(b): Let $P = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n]$ be the invertible matrix with ith column $\boldsymbol{p}_i$ such that $P^{-1}AP = D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix. Then $AP = PD$ and

$$\begin{aligned}[][A\boldsymbol{p}_1, A\boldsymbol{p}_2, \ldots, A\boldsymbol{p}_n] &= A[\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n] = AP = PD = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n]\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ &= [\lambda_1\boldsymbol{p}_1, \lambda_2\boldsymbol{p}_2, \ldots, \lambda_n\boldsymbol{p}_n]\end{aligned}$$

Hence $A\boldsymbol{p}_1 = \lambda_1\boldsymbol{p}_1$, $A\boldsymbol{p}_2 = \lambda_2\boldsymbol{p}_2$, $\ldots, A\boldsymbol{p}_n = \lambda_n\boldsymbol{p}_n$. Since P is invertible, $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ are linearly independent. Therefore A has n linearly independent eigenvectors.

(b)$\Rightarrow$(a): Suppose A has n linearly independent eigenvectors $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be corresponding eigenvalues, and $P = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n]$. Then P is invertible. Let $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then

$$AP = [A\boldsymbol{p}_1, A\boldsymbol{p}_2, \ldots, A\boldsymbol{p}_n] = [\lambda_1\boldsymbol{p}_1, \lambda_2\boldsymbol{p}_2, \ldots, \lambda_n\boldsymbol{p}_n] = PD.$$

Therefore $P^{-1}AP = D$ and A is diagonalizable. ∎

Theorem 2 (7.2.2) *If $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are eigenvectors of A corresponding to distinct elgenvalued $\lambda_1, \lambda_2, \ldots, \lambda_k$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is a linearly independent set.*

Proof. We may assume that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ is a linearly independent set and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{r+1}\}$ is a linearly dependent minimum set. Let

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_{r+1}\boldsymbol{v}_{r+1} = \boldsymbol{0}$$

Then by applying A on both hand sides, we have

$$c_1\lambda_1\boldsymbol{v}_1 + c_2\lambda_2\boldsymbol{v}_2 + \cdots + c_{r+1}\lambda_{r+1}\boldsymbol{v}_{r+1} = \boldsymbol{0}$$

Hence we have

$$c_1(\lambda_1 - \lambda_{r+1})\boldsymbol{v}_1 + \cdots + c_r(\lambda_r - \lambda_{r+1})\boldsymbol{v}_r = \boldsymbol{0}.$$

By our assumption, $c_1 = c_2 = \cdots = c_r = 0$. Thus $c_{r+1}\boldsymbol{v}_{r+1} = \boldsymbol{0}$ and $c_{r+1} = 0$ as $\boldsymbol{v}_{r+1} \neq \boldsymbol{0}$. This is a contradiction. ∎

Corollary 3 (7.2.3) *If an $n \times n$ matrix A has n distinct eigenvectors, then A is diagonalizable.*

Example 1

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{2,2} & a_{2,3} \\ 0 & 0 & a_{3,3} \end{bmatrix}$$

If $a_{1,1}, a_{2,2}, a_{3,3}$ are distinct, A is diagonalizable.

Example 2

$$A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

Then the characteristic polynomial $\det(tI - A)$ of A is $(t - \lambda)^3$. Hence eigenvalue is λ.

$$(\lambda I - A)\boldsymbol{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Then $\boldsymbol{x} = [t, 0, 0]^T$ and every eigenvector of A is a multiple of $\boldsymbol{e}_1$. Hence A is not diagonalizable.

Definition 2 If λ_0 is an eigenvalue of an $n \times n$ matrix A, then the dimension of the eigenspace corresponding to λ_0 is called the *geometric multiplicity* of λ_0 and the number of $\lambda - \lambda_0$ appears as a factor in the characteristic polynomial of A is called the *algebraic multiplicity* of A.

Theorem 4 (7.2.4) *If A is a square matrix, then*

(a) *For every eigenvalue of A, the geometric multiplicity is less than or equal to the algebraic multiplicity;*

(b) *A is diagonalizable if and only if, for every eigenvalue, the geometric multiplicity is equal to the algebraic multiplicity.*

Solutions to Homework 4

7.1.12 *Find the eigenvalues and bases for the eigenspaces of A^{25} for*

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

Solution.

$$\mathrm{char}_A(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda+1 & 2 & 2 \\ -1 & \lambda-2 & -1 \\ 1 & 1 & \lambda \end{vmatrix} = \lambda^3 - \lambda^2 - \lambda + 1 = (\lambda-1)^2(\lambda+1).$$

Let

$$\boldsymbol{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix},\ \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix},\ \boldsymbol{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix},\ \text{and } T = \begin{bmatrix} -1 & -1 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a bases for the eigenspaces of A and A^{25}. Moreover,

$$T^{-1}AT = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = D,\ T^{-1}A^{25}T = D,\ A^{25} = TDT^{-1} = A.$$

■

Note.

1. A の固有多項式 (characteristic polynomial) をここでは、$\mathrm{char}_A(\lambda)$ と書きました。この教科書では、決まった記号は使っていないようです。λ は変数ですから、他の変数 x や t などを使うこともできます。教科書にあるようにここでは、$\mathrm{char}_A(\lambda) = |\lambda I - A|$ を使いました。λ_0 が固有値であるとは、$\boldsymbol{v} \neq \boldsymbol{0}$ で $A\boldsymbol{v} = \lambda_0\boldsymbol{v}$ となる固有ベクトル $\boldsymbol{v}$ があることを言うのでした。この条件は、$(\lambda_0 I - A)\boldsymbol{v} = \boldsymbol{0}$ とも書き直せますから、

   λ_0 が固有値であるとは $(\lambda_0 I - A)\boldsymbol{x} = \boldsymbol{0}$ が非自明な解 [1]を持つことと同値

   でした。このことは、$\lambda_0 I - A$ が可逆ではない (not invertible)　と同値でしたから、$\det(\lambda_0 I - A) = 0$。すなわち、

   λ_0 が固有値であることと、$\det(\lambda_0 I - A) = 0$ とは同値

   でした。つまり、固有値を求めるには、固有多項式 $\mathrm{char}_A(\lambda) = |\lambda I - A|$ の解を求めれば良いことになります。同じようにして、$|A - \lambda I|$ の解を求めても構いません。実際 $A - \lambda I = -(\lambda I - A)$ ですから、各行に -1 をかける基本変形を考えれば、$|A - \lambda I| = (-1)^n|\lambda I - A|$ であることが分かりますから、解は同じです。それぞれ利点があり、$|\lambda I - A|$ は最高次の係数が 1 で、多項式としては扱いやすくなっています。一方、$-A$ を使うので、$\lambda I - A$ を書くときに注意しないといけません。逆に $|A - \lambda I|$ は行列式を書くのは簡単ですが、多項式のたとえば因数分解の時にまちがいを犯しやすいですね。いずれも注意が必要です。理論的には、$\mathrm{char}_A(\lambda) = |\lambda I - A|$ の方が扱いやすいので、ここではこちらを採用します。

2. 上の説明でも明らかなように、たとえば実数の範囲で考えていれば $\mathrm{char}_A(\lambda) = |\lambda I - A|$ の一つの（実数）根を λ_0 とすれば、$A\boldsymbol{v} = \lambda_0\boldsymbol{v}$ となる実ベクトル $\boldsymbol{v} \neq \boldsymbol{0}$ は必ず一つは見つかります。もし見つからなければ、それは、何かの計算まちがいです。数学ではよくまちがいを犯しますが、同時にまちがいを発見できる分野でもあります。理論を理解して、まちがいを早く発見しましょう。

[1] non-trivial solution, i.e., この方程式（斉次方程式 (homogeneous equation) とよびました）では $\boldsymbol{x} = \boldsymbol{0}$ は常に解でしたから、$\boldsymbol{x} \neq \boldsymbol{0}$ である解の存在が重要でそれを非自明な解と呼ぶのでした。

3. この問題では、固有値は 1 と -1 でした。固有ベクトルからなる（正確には固有ベクトルと $\mathbf{0}$ からなる）部分空間（固有ベクトル空間 eigenspace）$V(\lambda_0) = \{\boldsymbol{v} \mid (\lambda_0 I - A)\boldsymbol{v} = \mathbf{0}\}$ ですから $B = \lambda_0 I - A$ と考えれば、$B\boldsymbol{x} = \mathbf{0}$ の解を求めれば良いことになります。上で注意したように、λ_0 が固有値であれば、$\dim V(\lambda_0) \geq 1$ です。つまり解は無限個存在するので、解はその次元と同じ個数のパラメター表示によって表されるはずです。このときに必要なパラメターの数が、$\dim V(\lambda_0)$ ということになります。$V(\lambda_0)$ は $\lambda_0 I - A$ の Nullspace でしたから、その次元（nullity と言いました）は、n を行列 A のサイズとすると、$n - \mathrm{rank}(\lambda_0 I - A)$ となるのでした。ですから、次元だけ求めるには、$\lambda_0 I - A$ の階数、$\mathrm{rank}(\lambda_0 I - A)$ を求めれば十分です。しかし、固有空間の基底を求めようと思えば、実際に解を求める必要があります。

4. $\lambda_0 = 1$ のときは、

$$\lambda_0 I - A = \begin{bmatrix} 1+1 & 2 & 2 \\ -1 & 1-2 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

ですから、形から明らかに $\mathrm{rank}(\lambda_0 I - A) = 1$ で、解空間は、

$$V(1) = \left\{ s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \;\middle|\; s, t \in \boldsymbol{R} \right\} = \{s\boldsymbol{v}_1 + t\boldsymbol{v}_2 \mid s, t \in \boldsymbol{R}\}.$$

この基底は、解答に書いた $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ となります。上の計算で本来は拡大係数行列を用いますが、斉次で $\boldsymbol{b}$ に対応する最後の列は $\mathbf{0}$ なので単に係数行列の部分だけを書いています。

5. $\lambda_0 = 1$ のときは、

$$\lambda_0 I - A = \begin{vmatrix} -1+1 & 2 & 2 \\ -1 & -1-2 & -1 \\ 1 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 0 & 2 & 2 \\ -1 & -3 & -1 \\ 1 & 1 & -1 \end{vmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

ですから、 $\mathrm{rank}(\lambda_0 I - A) = 2$ で、解空間は、

$$V(-1) = \left\{ u \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \;\middle|\; u \in \boldsymbol{R} \right\} = \{u\boldsymbol{v}_3 \mid u \in \boldsymbol{R}\}.$$

この基底は、解答に書いた $\{\boldsymbol{v}_3\}$ となります。

6. これにより、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ が A の固有ベクトルによる基底であることが分かります。すなわち、対角化可能です。実は、これ自身が、 A^{25} の固有ベクトルによる基底になっています。一般に $A\boldsymbol{v} = \lambda_0 \boldsymbol{v}$ となっていれば、$AA\boldsymbol{v} = \lambda_0 A\boldsymbol{v} = \lambda_0^2 \boldsymbol{v}$ となり、これを続けると、$A^n \boldsymbol{v} = \lambda_0^n \boldsymbol{v}$ となるからです。固有値は、λ_0 から λ_0^{25} になりますが、固有ベクトルであることはかわりません。

7. ここで問題としてはおしまいですが、対角化との関連をみるために、$T = [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3]$ と置きました。すると、$\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ がそれぞれ固有値 1, 1, -1 に対応する固有ベクトルであることから、$A\boldsymbol{v}_1 = \boldsymbol{v}_1$, $A\boldsymbol{v}_2 = \boldsymbol{v}_2$, $A\boldsymbol{v}_3 = -\boldsymbol{v}_3$ となります。このことから、

$$AT = A[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3] = [A\boldsymbol{v}_1, A\boldsymbol{v}_2, A\boldsymbol{v}_3] = [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = TD, \text{ ただし } D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

となります。T の列は一次独立でしたから、T は可逆で $T^{-1}AT = D$ となります。ここで $(T^{-1}AT)^{25} = D^{25} = D$ であることに注意すると（$T^{-1}ATT^{-1}AT \cdots T^{-1}AT$ の間の $T^{-1}T$ は I となることに注意して下さい。また、D は対角線が 1 か -1 ですから 奇数乗である、25 乗してもかわりません。）これより $A^{25} = TDT^{-1}$ となりますが、最初の式から $TDT^{-1} = A$ でしたから $A^{25} = TDT^{-1} = A$ が分かりました。これは、固有値が 1, -1 と特殊でしたが、他の対角化可能な行列についても基本的には同じようにして、行列のべき（何乗）が求められます。

7.2.2 *Let*

$$A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}.$$

(a) *Find the eigenvalues of A.*

(b) *Find each eigenvalue λ, find the rank of the matrix $\lambda I - A$.*

(c) *Is A diagonalizable? Justify your conclusion.*

Solution.

(a) $\mathrm{char}_A(\lambda) = (\lambda - 5)(\lambda - 3)^2$. Hence $\lambda = 3$ and 5 are the eigenvalues.

(b) $\mathrm{rank}(3I - A) = 1$ and $\mathrm{rank}(5I - A) = 2$.

(c) A diagonalizable. By (b) the dimension of the eigenspace of 3 is two and 5 one. Thus A has a basis consisting of eigenvectors.

■

7.2.22 *Let*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Show that:

(a) *A is diagonalizable if $(a-d)^2 + 4bc > 0$.*

(b) *A is not diagonalizable if $(a-d)^2 + 4bc < 0$.*

Solution.

(a) If $(a-d)^2 + 4bc > 0$, the characteristic polynomial of A has two distinct real roots. Hence A is diagonalizable.

(b) If $(a-d)^2 + 4bc < 0$, the characteristic polynomial of A has no real roots. Hence A is not diagonalizable (as a real matrix).

■

Note.

1. 複素数の範囲で考えると、$(a-d)^2+4bc<0$ のときは、相異なる虚根を持つことになりますから、それぞれに対応して、複素ベクトルで固有ベクトルとなるものが見つかります。したがって、複素数の範囲では対角化可能です。一番簡単な例としては、$a=d=0$, $b=1$, $c=-1$ の場合を考えてみて下さい。固有多項式は λ^2+1 となります。固有値は $\sqrt{-1}$ と $-\sqrt{-1}$ です。固有ベクトルも求めることができますか。

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ \sqrt{-1} \end{bmatrix},\ \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ -\sqrt{-1} \end{bmatrix},\ U = [\boldsymbol{u}_1, \boldsymbol{u}_2] = \begin{bmatrix} 1 & 1 \\ \sqrt{-1} & -\sqrt{-1} \end{bmatrix},\ D = \begin{bmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{bmatrix}.$$

実際に $\boldsymbol{u}_1, \boldsymbol{u}_2$ が固有ベクトルであること、$AU = UD$ かつ $U^{-1}AU = D$ であることを確かめて下さい。

2. $(a-d)^2+4bc=0$ の場合はどうでしょうか。重根をもつので、これだけでは判定できません。しかし、対角化可能だとすると、重根を λ_0 として、それが、対角線に並ぶはずです。ということは、対角化したものは、$U^{-1}AU=\lambda_0 I$ の形をしています。すると $A=\lambda_0 UIU^{-1}=\lambda_0 I$ です。すなわち、最初から対角化されていることが分かります。したがって、$(a-d)^2+4bc=0$ でかつ、$\lambda_0 I$ の形をしていない場合は、対角化できません。無論、これは、サイズが 2 であることの特殊性です。2×2 の行列でない場合はずっと複雑です。

3. 対角化可能 $\Leftrightarrow$ 固有ベクトルによる基底を持つ $\Leftrightarrow$ 各固有値について geometric multiplicity = algebraic multiplicity。固有多項式の根がすべて異なれば対角化可能ですが、対角化可能だからといって固有多項式の根がすべて異なるわけではありません。

Solutions to Homework 5

7.3.3 *Find a matrix P that orthogonally diagonalizes A, and determine $P^{-1}AP$.*

$$A = \begin{bmatrix} 6 & 2\sqrt{3} \\ 2\sqrt{3} & 7 \end{bmatrix}.$$

Solution.

$$\mathrm{char}_A(\lambda) = \begin{vmatrix} \lambda - 6 & -2\sqrt{3} \\ -2\sqrt{3} & \lambda - 7 \end{vmatrix} = (\lambda-6)(\lambda-7) - 12 = \lambda^2 - 13\lambda + 30 = (\lambda-3)(\lambda-10).$$

For $\lambda = 3$,

$$\begin{bmatrix} 3-6 & -2\sqrt{3} \\ -2\sqrt{3} & 3-7 \end{bmatrix} \to \begin{bmatrix} 1 & 2/\sqrt{3} \\ 0 & 0 \end{bmatrix},\ \boldsymbol{v}_1 = \begin{bmatrix} -2/\sqrt{3} \\ 1 \end{bmatrix} \overset{\text{normalization}}{\longrightarrow} \boldsymbol{p}_1 = \frac{1}{\sqrt{7}}\begin{bmatrix} -2 \\ \sqrt{3} \end{bmatrix}.$$

For $\lambda = 10$,

$$\begin{bmatrix} 10-6 & -2\sqrt{3} \\ -2\sqrt{3} & 10-7 \end{bmatrix} \to \begin{bmatrix} 1 & -\sqrt{3}/2 \\ 0 & 0 \end{bmatrix},\ \boldsymbol{v}_2 = \begin{bmatrix} \sqrt{3}/2 \\ 1 \end{bmatrix} \overset{\text{normalization}}{\longrightarrow} \boldsymbol{p}_2 = \frac{1}{\sqrt{7}}\begin{bmatrix} \sqrt{3} \\ 2 \end{bmatrix}.$$

Hence

$$P = [\boldsymbol{p}_1, \boldsymbol{p}_2] = \frac{1}{\sqrt{7}}\begin{bmatrix} -2 & \sqrt{3} \\ \sqrt{3} & 2 \end{bmatrix},\ \text{ and } P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & 10 \end{bmatrix}. \quad \blacksquare$$

Note.

1. 実対称行列 A は実直交行列で対角化されます。すなわち、$P^TAP = D$ (D は対角行列) となる実直交行列 P が存在します。$A^T = A$ となる行列を対称行列 (symmetric matrix) と言います。そのうちで成分が実数のものが、実対称行列です。 $P^TP = I = PP^T$ となる行列を直交行列と言います。定義から P の列同志、行同志は互いに Euclid 内積に関して直交しかつ長さも 1 になっています。逆に、正規直交基底を列としてならべたものは直交行列になっています。

2. A を対称行列、$A\boldsymbol{u} = \lambda\boldsymbol{u}$, $A\boldsymbol{v} = \mu\boldsymbol{v}$ とします。すると $\boldsymbol{u}^TA = \boldsymbol{u}^TA^T = (A\boldsymbol{u})^T = \lambda\boldsymbol{u}^T$ であることに注意すると、
$$\lambda\boldsymbol{u}^T\boldsymbol{v} = \boldsymbol{u}^TA\boldsymbol{v} = \boldsymbol{u}^T(\mu\boldsymbol{v}) = \mu\boldsymbol{u}^T\boldsymbol{v}$$
となります。$\boldsymbol{u}^T\boldsymbol{v}$ は 1×1 行列ともスカラーとも見ることができますから、$\lambda \neq \mu$ ならば $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ であることが分かります。すなわち、対称行列の相異なる固有値に対する固有ベクトルは直交することが分かります。

3. 上で述べたことから、$n \times n$ 実対称行列が n 個の相異なる固有値 $\lambda_1, \lambda_2, \ldots, \lambda_n$ を持てば、それに対応する固有ベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ (但し $A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1, A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n$) は基底になりますが、これは直交基底であることが分かります。これから直交行列を作るには、これらを正規化 (normalize) して $\boldsymbol{p}_i = \frac{1}{\|\boldsymbol{v}_i\|}\boldsymbol{v}_i$ $(i = 1, 2, \ldots, n)$ とすれば良いことが分かります。（正規化はいずれにせよ最初のベクトルのスカラー倍になるのですから、正規化するベクトルのスカラー倍で簡単なものを取り、そのノルムを計算してそれで割った方が簡単ですね。このような事も、計算ミスを少なくする工夫です。）すなわち $\{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ は固有ベクトルからなる正規直交基底で、$P = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n]$ は直交行列で $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ で対角線に $\lambda_1, \lambda_2, \ldots, \lambda_n$ が並ぶ対角行列を表すとすると、
$$AP = [A\boldsymbol{p}_1, A\boldsymbol{p}_2, \ldots, A\boldsymbol{p}_n] = [\lambda_1\boldsymbol{p}_1, \lambda_2\boldsymbol{p}_2, \ldots, \lambda_n\boldsymbol{p}_n] = P\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$
より $P^TAP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ となります。

4. では固有多項式が重根を持つときはどうなるでしょうか。相異なる固有値に対する固有ベクトルは上で述べたように直交するのですが、同じ固有値に対する固有ベクトルはたとえ一次独立でも直交するとは限りません。そのときは、それぞれの固有空間の基底を Gram-Schmidt などを使って対角化する必要があります。

8.5.12 *In each part, find a basis for* $\boldsymbol{R}^3$ *relative to which the matrix for* T *is diagonal. In the following* $V(\lambda)$ *denote the eigenspace of* T *for* λ.

(a) $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -2x_1 + x_2 - x_3 \\ x_1 - 2x_2 - x_3 \\ -x_1 - x_2 - 2x_3 \end{bmatrix}$.

(b) $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -x_2 + x_3 \\ -x_1 + x_3 \\ x_1 + x_2 \end{bmatrix}$.

(c) $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 4x_1 + x_3 \\ 2x_1 + 3x_2 + 2x_3 \\ x_1 + 4x_3 \end{bmatrix}$.

Solution. Let $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be the standard basis and B' a basis relative to which the matrix of T is diagonal.

(a) Since $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -2x_1 + x_2 - x_3 \\ x_1 - 2x_2 - x_3 \\ -x_1 - x_2 - 2x_3 \end{bmatrix}$,

by computing $T(\boldsymbol{e}_1)$, $T(\boldsymbol{e}_2)$, $T(\boldsymbol{e}_3)$ or more precisely $[T(\boldsymbol{e}_1)]_B$, $[T(\boldsymbol{e}_2)]_B$, $[T(\boldsymbol{e}_3)]_B$ we find

$$[T]_B = [[T(\boldsymbol{e}_1)]_B, [T(\boldsymbol{e}_2)]_B, [T(\boldsymbol{e}_3)]_B] = \begin{bmatrix} -2 & 1 & -1 \\ 1 & -2 & -1 \\ -1 & -1 & -2 \end{bmatrix}.$$

Since $\mathrm{char}_T(\lambda)$ is independent of the choice of the basis, $\mathrm{char}_T(\lambda) = \lambda(\lambda+3)^2$.

$$V(0) = \mathrm{Span}\left(\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\right),\ V(-3) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right),\ \text{and } B' = \left\{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}.$$

(b) $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} -x_2 + x_3 \\ -x_1 + x_3 \\ x_1 + x_2 \end{bmatrix}$. Hence, $[T]_B = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Now $\mathrm{char}_T(\lambda) = (\lambda-1)^2(\lambda+2)$.

$$V(1) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right),\ V(-1) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right),$$

and

$$B' = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right\}.$$

(c) $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 4x_1 + x_3 \\ 2x_1 + 3x_2 + 2x_3 \\ x_1 + 4x_3 \end{bmatrix}$. Hence, $[T]_B = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$. Now $\mathrm{char}_T(\lambda) = (\lambda-5)(\lambda-3)^2$.

$$V(5) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right),\ V(3) = \mathrm{Span}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right),\ \text{and } B' = \left\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}.$$

Note.

1. この三つの一次変換で標準基底 (standard basis) に関して対称行列となっているのは、見て分かるように (a) と (b) のものです。ですからこれらを対角化する行列で直交行列となるものもあります。(a) については、$V(0)$ のベクトルと $V(3)$ のベクトルはすべて直交しています。これは、上の問題で述べたとおりです。しかし、B' の二つめと三つ目はともに、$V(3)$ のベクトルでかつ一次独立ですが、直交はしていません。この二つのベクトルを $\boldsymbol{v}_2, \boldsymbol{v}_3$ とすると、$\boldsymbol{v}_3' = [-1, 2, 1]^T = -\boldsymbol{v}_2 + 2\boldsymbol{v}_3$ となりこれは、$\boldsymbol{v}_2$ と直交します。 あとは、正規化して並べたものを P とすれば直交行列になります。ひと手間余分にかかりますが、直交行列は内積を変えないので、距離も変えず、とても重要なものです。最後の形を書いておきましょう。

$$P = \begin{bmatrix} \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix},\ P^{-1}[T]_B P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

   B' から作られる行列を P としても (P は直交行列ではありませんが) 同じ対角行列が現れます。対角線に並ぶのは、B' の固有ベクトルに対応するる固有値が順に並んでいることを確認して下さい。

2. 演習でも示しましたが、A 固有多項式は $T^{-1}AT$ の形の行列の固有多項式と同じになります。ですから、固有値も同じ、その重複度も同じです。教科書では p.434 と p.441 の Discussion Discovery に書いてあります。以下の式を見て下さい。

$$\begin{aligned} \mathrm{char}_{T^{-1}AT}(\lambda) &= \det(\lambda I - T^{-1}AT) = \det(\lambda T^{-1}IT - T^{-1}AT) \\ &= \det(T^{-1}(\lambda I - A)T) = \det(T^{-1})\det(\lambda I - A)\det(T) \\ &= \det(T^{-1})\det(T)\det(\lambda I - A) = \det(I)\det(\lambda I - A) = \det(\lambda I - A) \\ &= \mathrm{char}_A(\lambda). \end{aligned}$$

   ですから特に $T^{-1}AT$ が対角行列になっていれば、対角線に並ぶのは A の固有値ということになります。だから、固有ベクトルからなく基底が見つかり T が求まれば、 $T^{-1}AT$ を実際に計算しなくても、どのような対角行列になるのか、分かるわけです。

3. 今回の問題にはありませんが、J をすべての成分が 1 であるような、n 次正方行列としたとき対角線がすべて a で対角線以外がすべて b である $A = (a-b)I + bJ$ の形の行列については、J の対角化について理解していると簡単だという話をしました。J を対角化する直交行列に関して、少し間違った式も書いてしまいましたので、ここでまとめておきます。

   (a) J はまずどの行の成分を足しても n になりますからすべてが 1 である n 次列ベクトルを $\boldsymbol{j}$ とすると $J\boldsymbol{j} = n\boldsymbol{j}$ となります。これは、n が J の固有値で $\boldsymbol{v}_1 = \boldsymbol{j}$ が固有ベクトルであることを表しています。さて他の固有値はどうでしょうか。J は階数が 1 ですから Nullspace の次元は $n-1$ です。すなわち、Nullspace の基底となるようなベクトル $\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を取ると、$J\boldsymbol{v}_2 = J\boldsymbol{v}_3 = \cdots = J\boldsymbol{v}_n = \boldsymbol{0}$ となります。例えば、$\boldsymbol{v}_i = \boldsymbol{e}_1 - \boldsymbol{e}_i$ とすれば、良いことも分かります。ですから、ここで $T = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ とすると、

$$\begin{aligned} JT &= J[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] = [J\boldsymbol{v}_1, J\boldsymbol{v}_2, \ldots, J\boldsymbol{v}_n] = [n\boldsymbol{v}_1, \boldsymbol{0}, \ldots, \boldsymbol{0}] \\ &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]\mathrm{diag}(n, 0, \ldots, 0) = T\mathrm{diag}(n, 0, \ldots, 0). \end{aligned}$$

   となります。ここで $D = \mathrm{diag}(n, 0, \ldots, 0)$ と置けば、$T^{-1}JT = D$ となります。まとめると、J は 固有値 n を持ちその重複度は 1、それ以外に固有値 0 を重複度 $n-1$ で持つことが分かります。$\mathrm{char}_J(\lambda) = (\lambda - n)\lambda^{n-1}$ となります。

   (b) $A = (a-b)I + bJ$ とします。上の T を使って $T^{-1}AT$ を計算してみると

$$\begin{aligned} T^{-1}AT &= T^{-1}((a-b)I + bJ)T = (a-b)T^{-1}IT + bT^{-1}JT = (a-b)I + bD \\ &= \mathrm{diag}(a + b(n-1), a-b, \ldots, a-b). \end{aligned}$$

   となります。すなわち、A は J と同じ固有ベクトルを持ち、固有値は $a+b(n-1)$ を重複度 1 で、$a-b$ を重複度 $n-1$ で持つことになります。特に $\mathrm{char}_A(\lambda) = (\lambda - a - b(n-1))(\lambda - a + b)^{n-1}$ となります。一般に求まってしまいました。

(c) A は実は対称行列になっています。無論 J もそうです。ということは、単に対角化ではなく、直交行列で対角化できるはずです。その直交行列はどうなるでしょうか。$\boldsymbol{v}_1$ とそれ以外の $\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ は直交しますが、$\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ はこのままでは直交しません。実はこれに関してもいろいろな面白い話があるのですが、ここでは、上で使った T と、一つだけ、A を対角化する直交行列を書いておくことにします。みなさんも考えてみて下さい。

$$T = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & -1 & \cdots & 0 \\ & \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} & \\ 1 & 0 & 0 & 0 & \cdots & -1 \end{bmatrix}.$$

$$P = \begin{bmatrix} 1/\sqrt{n} & 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & \cdots & 1/\sqrt{(n-1)n} \\ 1/\sqrt{n} & -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & \cdots & 1/\sqrt{(n-1)n} \\ 1/\sqrt{n} & 0 & -2 & 1/\sqrt{12} & \cdots & 1/\sqrt{(n-1)n} \\ 1/\sqrt{n} & 0 & 0 & -3/\sqrt{12} & \cdots & 1/\sqrt{(n-1)n} \\ & \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} & \\ 1/\sqrt{n} & 0 & 0 & 0 & \cdots & (-n+1)/\sqrt{(n-1)n} \end{bmatrix}.$$

このように並べてみると、P の列は T の列 $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を Gram-Schidt で直交化したものですね。

Solutions to Homework 6

10.6.12 *Find a unitary matrix P that diagonalizes A, and determine $P^{-1}AP$.*

$$A = \begin{bmatrix} 2 & \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & 2 & 0 \\ \frac{i}{\sqrt{2}} & 0 & 2 \end{bmatrix}$$

Solution.

$$\mathrm{char}_A(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - 2 & -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \lambda - 2 & 0 \\ -\frac{i}{\sqrt{2}} & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 2)(\lambda - 3).$$

$$P = \begin{bmatrix} \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \end{bmatrix}, \text{ and } P^{-1} = P^* = \begin{bmatrix} -\frac{i}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}. \text{ Hence}$$

$$P^*AP = \begin{bmatrix} -\frac{i}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & \frac{i}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ -\frac{i}{\sqrt{2}} & 2 & 0 \\ \frac{i}{\sqrt{2}} & 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{i}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

■

10.6.23 *Let λ and μ be distinct eigenvalues of a Hermitian matrix A.*

(a) *Prove that if $\boldsymbol{x}$ is an eigenvector corresponding to λ and $\boldsymbol{y}$ an eigenvector corresponding to μ, then $\boldsymbol{x}^*A\boldsymbol{y} = \lambda\boldsymbol{x}^*\boldsymbol{y}$ and $\boldsymbol{x}^*A\boldsymbol{y} = \mu\boldsymbol{x}^*\boldsymbol{y}$.*

(b) *Show that $\boldsymbol{x}^*\boldsymbol{y} = 0$. Hence $\boldsymbol{x}\cdot\boldsymbol{y} = \overline{\boldsymbol{x}^*\boldsymbol{y}} = 0$.*

(b') *If A is a normal matrix, then eigenvalues from different eigenspaces of A are orthogonal.*

Solution.

(a) First note that by Theorem 10.6.5, the eigenvalues of a Hermitian matrix are real numbers. Since $A\boldsymbol{x} = \lambda\boldsymbol{x}$, $\lambda\boldsymbol{x}^* = \overline{\lambda}\boldsymbol{x}^* = (A\boldsymbol{x})^* = \boldsymbol{x}^*A^*$. Hence $\mu\boldsymbol{x}^*\boldsymbol{y} = \boldsymbol{x}^*(\mu\boldsymbol{y}) = \boldsymbol{x}^*A\boldsymbol{y} = \lambda\boldsymbol{x}^*\boldsymbol{y}$ ■

(b) By (a), $(\lambda - \mu)\boldsymbol{x}^*\boldsymbol{y} = 0$. Since λ and μ are distinct, $\boldsymbol{x}^*\boldsymbol{y} = 0$. ■

(b') Let $A\boldsymbol{x} = \lambda\boldsymbol{x}$ for some nonzero vector $\boldsymbol{x}$. Since A is normal, $A^T\overline{A} = \overline{A^*A} = \overline{AA^*} = \overline{A}A^T$. Hence

$$\begin{aligned}
0 &= \|A\boldsymbol{x} - \lambda\boldsymbol{x}\|^2 = (A\boldsymbol{x} - \lambda\boldsymbol{x})^T\overline{(A\boldsymbol{x} - \lambda\boldsymbol{x})} = (\boldsymbol{x}^TA^T - \lambda\boldsymbol{x}^T)(\overline{A}\overline{\boldsymbol{x}} - \overline{\lambda}\overline{\boldsymbol{x}}) \\
&= \boldsymbol{x}^TA^T\overline{A}\overline{\boldsymbol{x}} - \overline{\lambda}\boldsymbol{x}^TA^T\overline{\boldsymbol{x}} - \lambda\boldsymbol{x}^T\overline{A}\overline{\boldsymbol{x}} + |\lambda|^2\boldsymbol{x}^T\overline{\boldsymbol{x}} \\
&= \boldsymbol{x}^T\overline{A}A^T\overline{\boldsymbol{x}} - \overline{\lambda}\boldsymbol{x}^TA^T\overline{\boldsymbol{x}} - \lambda\boldsymbol{x}^T\overline{A}\overline{\boldsymbol{x}} + |\lambda|^2\boldsymbol{x}^T\overline{\boldsymbol{x}} \\
&= \boldsymbol{x}^TA^{*T}A^T\overline{\boldsymbol{x}} - \overline{\lambda}\boldsymbol{x}^TA^T\overline{\boldsymbol{x}} - \lambda\boldsymbol{x}^TA^{*T}\overline{\boldsymbol{x}} + |\lambda|^2\boldsymbol{x}^T\overline{\boldsymbol{x}} \\
&= (\boldsymbol{x}^TA^{*T} - \overline{\lambda}\boldsymbol{x}^T)(A^T\overline{\boldsymbol{x}} - \lambda\overline{\boldsymbol{x}}) = (A^*\boldsymbol{x} - \overline{\lambda}\boldsymbol{x})^T\overline{(A^*\boldsymbol{x} - \overline{\lambda}\boldsymbol{x})} = \|A^*\boldsymbol{x} - \overline{\lambda}\boldsymbol{x}\|^2.
\end{aligned}$$

Thus $A^*\boldsymbol{x} = \overline{\lambda}\boldsymbol{x}$. Hence $\boldsymbol{x}^*A = \lambda\boldsymbol{x}^*$. Let $A\boldsymbol{y} = \mu\boldsymbol{y}$. Then

$$\mu\overline{\boldsymbol{x}^T\overline{\boldsymbol{y}}} = \mu\boldsymbol{x}^*\boldsymbol{y} = \boldsymbol{x}^*A\boldsymbol{y} = \lambda\boldsymbol{x}^*\boldsymbol{y} = \lambda\overline{\boldsymbol{x}^T\overline{\boldsymbol{y}}}.$$

Hence if $\lambda \neq \mu$, then $\boldsymbol{x}\cdot\boldsymbol{y} = \boldsymbol{x}^T\overline{\boldsymbol{y}} = 0$. ■

最初に A をエルミートとしてありますから、(b) が聞きたかった問題でしょう。エルミート行列 $(A^* = A)$ は、正規 $(AA^* = A^*A)$ だから (b) は (b') の特別な場合ということになります。

上で示したように一般に A が正規行列、$A\boldsymbol{v} = \lambda\boldsymbol{v}$ とすると、$A^*\boldsymbol{v} = \overline{\lambda}\boldsymbol{v}$ が成り立つ。

また一般に A を n 次正方行列とすると A^*A はエルミート行列となるから、A^*A の固有値は実数である。

Note.

1. 一般に、二つの正方行列 A, B が交換可能 $AB = BA$ とし、$V(\lambda)$ を A の固有値 λ に関する固有空間とします。すると、$\boldsymbol{v} \in V(\lambda)$ について $A\boldsymbol{v} = \lambda\boldsymbol{v}$ となります。しかし、

$$AB\boldsymbol{v} = BA\boldsymbol{v} = B(\lambda\boldsymbol{v}) = \lambda B\boldsymbol{v}$$

だから $B\boldsymbol{v} \in V(\lambda)$ すなわち、$BV(\lambda) \subseteq V(\lambda)$ となります。

$$T_B : V(\lambda) \to V(\lambda)\ (\boldsymbol{v} \mapsto B\boldsymbol{v})$$

は線形写像 (linear mapping ですから、T_B の $V(\lambda)$ における固有値を μ, 固有ベクトルを $\boldsymbol{v}$ とすると、$\boldsymbol{v} \in V(\lambda)$ ですから、$A\boldsymbol{v} = \lambda\boldsymbol{v}$ かつ $B\boldsymbol{v} = \mu\boldsymbol{v}$ となり、$\boldsymbol{v}$ は A と B の共通の固有ベクトルとなります。特に、A も B も対角化可能、すなわち、固有ベクトルからなる基底を持つときは、上の議論を続けることにより、A と B の共通の固有ベクトルからなる基底を持つことが分かります。すなわち、

> A, B を可換な (i.e., $AB = BA$) それぞれが対角化可能な n 次正方行列とすると、A, B の共通の固有ベクトルからなる基底を持ち、その基底を列ベクトルとしてならべた行列を T とすると、$T^{-1}AT$ も $T^{-1}BT$ も対角行列になる。すなわち、同時対角化可能である。

2. $V = \boldsymbol{C}^n$ (または $V = \boldsymbol{R}^n$) を内積空間とし W をその部分空間 A を n 次正方行列とする。$AW \subseteq W$ すなわち、$A\boldsymbol{w} \in W$ が任意の $\boldsymbol{w} \in W$ について成立すれば、$A^*W^\perp \subseteq W^\perp$ である（$V = \boldsymbol{R}^n$ のときは $A^* = A^T$)。実際、$\boldsymbol{v} \in W^\perp$ とすると、任意の $\boldsymbol{w} \in W$ について $A\boldsymbol{w} \in W$ だったから、

$$A^*\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{v}^T\overline{A}\overline{\boldsymbol{w}} = \boldsymbol{v}^T\overline{A\boldsymbol{w}} = 0.$$

これは、$A^*\boldsymbol{v} \in W^\perp$ を表す。

3. A を正規行列とする。λ を A の固有値として、$V(\lambda)$ を A の固有値 λ に関する固有空間とする。まず 1 より $A^*V(\lambda) \subseteq V(\lambda)$。次に 2 より $AV(\lambda)^\perp \subseteq V(\lambda)^\perp$. $V = V(\lambda) + V(\lambda)^\perp$ かつ $V(\lambda) \cap V(\lambda)^\perp = \{\boldsymbol{0}\}$ であるが、ここで $A\boldsymbol{v} = \mu\boldsymbol{v}$、$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$, $\boldsymbol{u} \in V(\lambda)$, $\boldsymbol{w} \in V(\lambda)^\perp$ とすると、

$$\mu\boldsymbol{u} + \mu\boldsymbol{w} = \mu\boldsymbol{v} = A\boldsymbol{v} = A\boldsymbol{u} + A\boldsymbol{w} = \lambda\boldsymbol{u} + A\boldsymbol{w}$$

より $(\mu - \lambda)\boldsymbol{u} = A\boldsymbol{w} - \mu\boldsymbol{w} \in V(\lambda)^\perp$ で左辺は $V(\lambda)$ に入っているので $\boldsymbol{0}$ となる。これより、$\mu \neq \lambda$ の場合は $\boldsymbol{u} = \boldsymbol{0}$ となる。すなわち、$\boldsymbol{v} \in V(\lambda)^\perp$ となる。これは、λ とは異なる固有値に対する固有ベクトルは $V(\lambda)^\perp$ に入ることを意味するから、λ を固有値とするベクトルとは直交する。すなわち、(b') の別証明を与える。

Chapter 10. 10 *Suppose that $A^* = -A$.*

(a) *Show that iA is Hermitian.*

(b) *Show that A is unitary diagonalizable and has pure imaginary eigenvalues.*

Solution.

(a) Since $(iA)^* = \overline{iA^T} = -iA^* = iA$, iA is Hermitian. ∎

(b) Since iA is Hermitian, iA is diagonalizable by a unitary matrix U. Then $U^*(iA)U = D$ for some diagonal matrix D. Hence $U^*AU = -iD$. Since iA is Hermitian, eigenvalues of iA is real by Theorem 10.6.5. Hence all diagonal entries of D are real, and all eigenvalues of A which are the diagonal entries of $-iD$ are pure imaginary. ∎

Note. Since $A^* = -A$ and A commutes with $-A$, A is a normal matrix. Hence A is unitary diagonalizable. Suppose $A\boldsymbol{v} = \lambda\boldsymbol{v}$ for some nonzero vector $\boldsymbol{v}$. Then $\bar{\lambda}\boldsymbol{v}^* = \boldsymbol{v}^*A^*$ and

$$-\bar{\lambda}\|\boldsymbol{v}\|^2 = -\bar{\lambda}\boldsymbol{v}^*\boldsymbol{v} = -\boldsymbol{v}^*A^*\boldsymbol{v} = \boldsymbol{v}^*A\boldsymbol{v} = v^*\lambda\boldsymbol{v} = \lambda\|\boldsymbol{v}\|^2.$$

Therefore $\bar{\lambda} = -\lambda$ and λ is purely imaginary.

Solutions to Homework 7

7.2.25 *Indicate whether each statement is always true or sometimes false. Justify your answer by giving a logical argument or a counterexample.*

(a) *A square matrix with linearly independent column vectors is diagonalizable.*

(b) *If A is diagonalizable, then there is a unique matrix P such that $P^{-1}AP$ is a diagonal matrix.*

(c) *If $\boldsymbol{v}_1, \boldsymbol{v}_2$, and $\boldsymbol{v}_3$ come from different eigenspaces of A, then it is impossible to express $\boldsymbol{v}_3$ as a linear combination of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$.*

(d) *If A is diagonalizable and invertible, then A^{-1} is diagonalizable.*

(e) *If A is diagonalizable, then A^T is diagonalizable.*

Solution.

(a) False. The following square matrix has linearly independent column vectors with eigenvalue 1. So if it is diagonalizable it is similar to the identity matrix, which is impossible.

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \text{ and } T^{-1}AT \neq I \text{ as } A \neq I = TIT^{-1}$$

for any invertible matrix T. Since $\mathrm{rank}(I - A) = 1$, the dimension of eigenspace corresponding to eigenvalue 1 is 1. Hence A is not diagonaliable. ■

(b) False. If $T^{-1}AT = D$ is a diagonalization, T can be replaced by cT with $c \neq 0$ as $(cT)^{-1}A(cT) = T^{-1}AT = D$. ■

(c) True. It is because eigenvectors come from different eigenspaces are linearly independent. If $\boldsymbol{v}_3 = a\boldsymbol{v}_1 + b\boldsymbol{v}_2$, then $a\boldsymbol{v}_1 + b\boldsymbol{v}_2 - \boldsymbol{v}_3 = \boldsymbol{0}$. Hence $\boldsymbol{v}_3 = \boldsymbol{0}$ and $a = b = 0$. ■

(d) True. Let $T^{-1}AT = D$ be a diagonalization. Then $T^{-1}A^{-1}T = D^{-1}$. In particular, all eigenvalues of A are nonzero and A^{-1} is diagonalizable by the same matrix that diagonalized A. ■

(e) True. Let $T^{-1}AT = D$ be a diagonalization. Then $T^TA^T(T^{-1})^T = D^T = D$. Since $(T^{-1})^T = (T^T)^{-1}$ as $T^{-1}T = I = TT^{-1}$ implies $T^T(T^{-1})^T = I = (T^{-1})^TT^T$. In particular, all eigenvalues of A are same as that of A^T and A^T is diagonalizable by T^T if T diagonalizes A. ■

Note.

1. (a) は linearly independent column vectors という phrase に惑わされないようにという注意でしょう。$n \times n$ 行列 A が n 個の一次独立な、固有ベクトル (eigenvectors) すなわち、固有ベクトルからなる基底をもてば、それを、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, $A\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}$, $A\boldsymbol{v}_2 = \lambda_2\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n = \lambda_n\boldsymbol{v}_n$ とすれば、これらを列ベクトルとする $T = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$ とし、$D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ とすると、T を構成する列ベクトルが一次独立であることから T は可逆 (invertible) となり $T^{-1}AT = D$ となるのでした。この問題では、A の固有ベクトルではなく、A 自体の列ベクトルが一次独立という条件ですから、これから得られるのは、A が可逆だということだけです。A が 可逆なら、$A\boldsymbol{v} = \boldsymbol{0} \Rightarrow \boldsymbol{v} = \boldsymbol{0}$ ですから、0 は固有値ではありません。すなわち、可逆な行列の固有値はすべて 0 でないことが分かります。逆に固有値がすべて 0 でなければ、固有多項式 $\mathrm{char}_A(\lambda)$ は $\lambda = 0$ を根に持ちませんから、$0 \neq \mathrm{char}_A(0) = \det(0 \cdot I - A) = \det(-A) = (-1)^n \det(A)$ となります。すなわち、$\det(A) \neq 0$ で可逆となります。まとめると、

$$A \text{ の固有値は零ではない} \Leftrightarrow A \text{ は可逆} \Leftrightarrow A \text{ の列ベクトルは一次独立。}$$

2. (a) の反例では、固有値が 1 のサイズ 2 のジョルダンセルを使いましたが、固有値が零でない n 次のジョルダンセルは、すべて、一次独立な、列ベクトルをもち、かつ、$n \geq 2$ ならば対角化できません。$J(n, \lambda)$ でそのような行列を表すと

$$J(n,\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ 0 & \cdots & \cdots & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda \end{bmatrix}, \ \lambda I - J(n,\lambda) = \begin{bmatrix} 0 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ 0 & \cdots & \cdots & \ddots & 0 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{bmatrix}$$

だから $\mathrm{rank}(\lambda I - J(n, \lambda)) = n - 1$ となり、$J(n, \lambda)$ の固有値はすべて λ でその固有空間の次元は、1 となりますから、$n \geq 2$ ならば対角化可能ではありません。

3. (b) は対角化を仮定して、対角化する行列、この場合は P が一通りに決まるかという問題です。上の解答では定数倍を用いましたが、$P(i, j)$ を 基本行列で I の i 行と j 行を交換して作ったものとすると、$P(i, j)$ を左からかけることは i 行と j 行の入れ替え、右からかけることは、i 列と j 列の列の入れ替えとなり、かつ $P(i, j) = P(i, j)^{-1}$ であることは簡単に分かりますから、$P^{-1}AP = D$ とすると $P(i, j)^{-1}DP(i, j) = P(i, j)DP(i, j)$ は対角行列 D の i 番目と、j 番目を入れ替えたものになります。ですから $P' = PP(i, j)$ とすると $P'^{-1}AP'$ はやはり対角行列になります。これを様々な i, j に続けてかけてもよいので、D の対角成分の並び替えをしてよければ、その並び替えの分だけ、いろいろな P を取れることになります。では、$P(i; c)$ はどうでしょうか。ただし、$c \neq 0$ とします。$P'' = PP(i; c)$ としたときは、$P''^{-1}AP''$ はどうなるでしょうか。考えてみて下さい。もちろんその次は $P(i, j; c)$ ですよね。

4. $P^{-1}AP = D = P'^{-1}AP'$ となっている場合を考えましょう。すると $DP^{-1}P' = P^{-1}P'D$ となります。ですから D と交換可能な行列が分かっていれば、$P^{-1}P'$ はそのような行列ですから、これから P と P' はどのくらい異なるかが分かります。では $QD = DQ$ となる可逆行列をすべて求めることはできますか。なかなか良い問題ですよ。

5. (c) は問題無かったと思いますが、あまり Linear Algebra II や Advanced Linear Algebra の授業で説明していなかったと思われるベクトル空間の直和のことを書きます。

   V をベクトル空間、$W_1, W_2, \ldots, W_m$ をその部分空間とします。このとき、

$$W_1 + W_2 + \cdots + W_m = \{\boldsymbol{w}_1 + \boldsymbol{w}_2 + \cdots + \boldsymbol{w}_m \mid \boldsymbol{w}_1 \in W_1,\ \boldsymbol{w}_2 \in W_2, \ldots, \boldsymbol{w}_m \in W_m\}$$

   すなわち、それぞれの部分空間からベクトルをとって足したものとすると、これは、また、V の部分空間となります。これを、$W_1, W_2, \ldots, W_m$ の和といいます。さて、$V = W_1 + W_2 + \cdots + W_m$ としたとき、$\boldsymbol{v} = \boldsymbol{w}_1 + \boldsymbol{w}_2 + \cdots + \boldsymbol{w}_m$, $(\boldsymbol{w}_1 \in W_1,\ \boldsymbol{w}_2 \in W_2, \ldots, \boldsymbol{w}_m \in W_m)$ と表せますが、この表し方が一通りしかないとき、V を $W_1, W_2, \ldots, W_m$ の直和といい、$V = W_1 \oplus W_2 \oplus \cdots \oplus W_m$ とか $V = W_1 \dot{+} W_2 \dot{+} \cdots \dot{+} W_m$ と書きます。一通りということは、$\boldsymbol{v} = \boldsymbol{w}_1 + \boldsymbol{w}_2 + \cdots + \boldsymbol{w}_m = \boldsymbol{w}'_1 + \boldsymbol{w}'_2 + \cdots + \boldsymbol{w}'_m$ ならば $\boldsymbol{w}_1 = \boldsymbol{w}'_1,\ \boldsymbol{w}_2 = \boldsymbol{w}'_2, \ldots, \boldsymbol{w}_m = \boldsymbol{w}'_m$ であることを表します。また内積空間で、$i \neq j$ のとき $\boldsymbol{w}_i \cdot \boldsymbol{w}_j = 0$ が成り立つ場合は特に、直交直和といい、$V = W_1 \perp W_2 \perp \cdots \perp W_m$ と書いたりします。一通りにかけるというのは、何か、一次独立性と似ていると思います。実は、$\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m$ をゼロでないベクトルとして、$W_i = \mathrm{Span}(\boldsymbol{w}_i)$ とすると、$\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m$ が一次独立ということと、$W_1 + W_2 + \cdots + W_m$ が直和であることとが同値であることが分かります。つまり、直和は、一次独立ということを、ベクトルではなく、部分空間に拡張したものになっています。この言葉を使って、基底を表現するとこのようになります。V を n 次元ベクトル空間とするとき、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ が基底であることと、$V = \mathrm{Span}(\boldsymbol{v}_1) \oplus \mathrm{Span}(\boldsymbol{v}_2) \oplus \cdots \oplus \mathrm{Span}(\boldsymbol{v}_n)$ であることとは同値です。また、内積空間で、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ が直交基底であることと、$V = \mathrm{Span}(\boldsymbol{v}_1) \perp \mathrm{Span}(\boldsymbol{v}_2) \perp \cdots \perp \mathrm{Span}(\boldsymbol{v}_n)$ であることとは同値です。

6. さて、n 次正方行列 A の相異なる固有値を $\lambda_1, \lambda_2, \ldots, \lambda_r$ とし、固有多項式を $\mathrm{char}_A(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_r)^{m_r}$ とします。（実数ではこのように因数分解できない場合もありますが、複素数まで考えれば、このように必ず分解できることが知られています。）さて、そこで、$W_i =$

$V(\lambda_i) = \{\boldsymbol{v} \in V \mid A\boldsymbol{v} = \lambda_i \boldsymbol{v}\}$ を A の 固有値 λ_i に関する固有空間とすると、W_i は V の部分空間ですが、$W_1 + W_2 + \cdots + W_r$ は直和になっています。$1 \leq \dim W_i \leq m_i$ です。固有多項式の形から $n = m_1 + m_2 + \cdots + m_r$ ですから、次元を考えると、$\dim W_i = m_i$ がすべての i について成立することと、$V = W_1 + W_2 + \cdots + W_r$ となることと同値です。すなわち、λ_i の algebraic multiplicity m_i と geometric multiplicity $\dim W_i$ がすべての i について等しければ、$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ となり、固有ベクトルからなる、基底が取れるので、対角化が可能になります。

7. 通常は、$V \supset W_1 \oplus W_2 \oplus \cdots \oplus W_r$ となっているので、V を固有空間に変わるものの直和にかけないかと考えると、Jordan の標準型に至るのです。ここでは詳しく述べませんが、$\widetilde{W_i} = \{\boldsymbol{v} \mid (\lambda_i I - A)^{m_i} \boldsymbol{v} = \boldsymbol{0}\}$ とすると、$W_i \subset \widetilde{W_i}$ となり、かつ
$$V = \widetilde{W_1} \oplus \widetilde{W_2} \oplus \cdots \oplus \widetilde{W_r} \supset W_1 \oplus W_2 \oplus \cdots \oplus W_r$$
となっています。$\widetilde{W_i}$ を固有値 λ_i に関する一般固有空間 (generallzed eigenspace) と言います。

8. (d) (e) で分かったように、A が対角化可能ならば、A^T も対角化可能、また、A が可逆なときは、A^{-1} も対角化可能となっています。同様に考えると A^i $(i = 1, 2, \ldots)$ も対角化可能ですし、A が可逆なときは、A^i で i が負のときも対角化可能です。それは、$(T^{-1}AT)^i = T^{-1}A^iT$ などから明らかです。一つ注意しないといけないのは、A と A^i は同じ可逆行列 T で対角化できますが、A^T は A と同じ行列で対角化可能だとは限りません。同じ T で対角化可能なためには、A と A^T とが交換可能でないといけません。それは、対角化したあとは、交換可能ですから、明らかです。

7.2.26 *Suppose that the characteristic polynomial of some matrix A is found to be $p(\lambda) = (\lambda - 1)(\lambda - 3)^2(\lambda - 4)^3$. In each part, answer the question and explain your reasoning.*

(a) *What can you say about the dimensions of the eigenspaces of A?*

(b) *What can you say about the dimension of the eigenspaces if you know that A is diagonalizable?*

(c) *If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a linearly independent set of eigenvalues of A all of which correspond to the same eigenvalue of A, what can you say about the eigenvalue?*

Solution. Let $V(\lambda_0)$ denote the eigenspace of A corresponding to the eigenvalue λ_0.

(a) $\dim V(1) = 1$, $1 \leq \dim V(3) \leq 2$ and $1 \leq \dim V(4) \leq 3$. ■

(b) $\dim V(1) = 1$, $\dim V(3) = 2$ and $\dim V(4) = 3$. ■

(c) The assumption states that $\dim V(\lambda_0) \geq 3$. Hence by (a), $\lambda_0 = 4$ and these vectors belong to the eigenspace corresponding to the eigenvalue 4. ■

7.2.29 *Show that the Jordan block matrix J_n has $\lambda = 1$ as its only eigenvalue and that the corresponding eigenspace is* $\mathrm{Span}(\boldsymbol{e}_1\}$.

Solution. Clearly $\mathrm{char}_{J_n}(\lambda) = \det(\lambda I - J_n) = \det(\mathrm{diag}(\lambda - 1, \lambda - 1, \ldots, \lambda - 1)) = (\lambda - 1)^n$. Hence the only eigenvalue of J_n is 1. Clearly $J_n\boldsymbol{e}_1 = \boldsymbol{e}_1$ and $\mathrm{rank}(I - J_n) = n - 1$. Hence the eigenspace of J_n corresponding to 1 is of dimension 1 and equal to $\mathrm{Span}(\boldsymbol{e}_1)$. ■

Solutions to Homework 8

9.5.12 (a)(c)(d) *In each part, classify the matrix as positive definite, positive semidefinite, negative definite, negative semidefinite, or indefinite.*

$$(a)\begin{bmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad (c)\begin{bmatrix} 6 & 7 & 1 \\ 7 & 9 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \qquad (d)\begin{bmatrix} -4 & 7 & 8 \\ 7 & -3 & 9 \\ 8 & 9 & -1 \end{bmatrix}.$$

Solution. Let A be the corresponding matrix in each case.

(a) Since $\boldsymbol{e}_1^T A\boldsymbol{e}_1 = 3 > 0$ and $\boldsymbol{e}_2^T A\boldsymbol{e}_2 = -2 < 0$, A is indefinite.

(c) Since $\mathrm{char}_A(\lambda) = \lambda^3 - 16\lambda^2 + 15\lambda = \lambda(\lambda-1)(\lambda-15)$, eigenvalues are $0, 1$ and 15. Hence A is positive semidefinite.

(d) Since $\mathrm{char}_A(\lambda) = \lambda^3 + 8\lambda^2 - 175\lambda - 1561$, the product of three real roots of $\mathrm{char}_A(\lambda)$ is 175 and hence positive, and the sum of them is -8 and hence negative. Therefore three roots are one positive and two negatives. Hence it is indefinite. [Since $\mathrm{char}_A(-11) = 1 > 0$ and $\mathrm{char}_A(0) = -1561 < 0$, we can conclude the same as $\mathrm{char}_A(x) \to -\infty$ if $x \to -\infty$ and $\mathrm{char}_A(x) \to \infty$ if $x \to \infty$.] ■

Note.

1. A を n 次対称行列としたとき、A が正値または正定値 (positive definite) であるとは、任意の $\boldsymbol{0} \neq \boldsymbol{x} \in \boldsymbol{R}^n$ について $\boldsymbol{x}^T A\boldsymbol{x} > 0$ であることを言います。半正値 (positive semidefinite) は $\boldsymbol{x}^T A\boldsymbol{x} \geq 0$ などとなります。正にも負にもなるとき、indefinite (不定符号（この日本語は一般的ではない）) と呼ぶわけです。$\boldsymbol{x}^T A\boldsymbol{x} < 0$ であれば、$\boldsymbol{x}^T(-A)\boldsymbol{x} > 0$ ですから、基本的には正値または半正値の判定ができれば十分です。

2. 判定条件の前に基本的なこととして、$\boldsymbol{e}_i$ を標準基底、A の (i,i) 成分を $a_{i,i}$ とすると $\boldsymbol{e}_i^T A\boldsymbol{e}_i = a_{i,i}$ となります。計算すれば分かりますが、$A\boldsymbol{e}_i$ は A の第 i 列となりますから、それと $\boldsymbol{e}_i^T$ の積は $a_{i,i}$ となります。ですから、A が正値であれば、$0 < \boldsymbol{e}_i^T A\boldsymbol{e}_i = a_{i,i}$ です。つまり、(a) はこの事だけから、indefinite である (正値にも半正値にも、負値にも半負値でもない) ことがわかります。また (c) は半負値にはならない。(d) は半正値にはならないことが分かります。

3. 主小行列は、左上から i 行 i 列だけをとってつくった正方行列です。その行列式を主小行列式といいます。たとえば、(d) の行列の場合次の三つが主小行列式です。

$$|-4| = -4 \text{ 絶対値ではなく行列式}, \begin{vmatrix} -4 & 7 \\ 7 & -3 \end{vmatrix} = -37, \begin{vmatrix} -4 & 7 & 8 \\ 7 & -3 & 9 \\ 8 & 9 & -1 \end{vmatrix} = 1561.$$

4. A を n 次対称行列としたとき、以下は同値である。

   (i) A は正値

   (ii) A の固有値はすべて正

   (iii) A の主小行列式は正

   Proof. (i) ⇔ (ii) は下の Theorem 9.5.1 およびその証明から明らか。$\|\boldsymbol{x}\| = 1$ の仮定があるが、証明では、$\boldsymbol{x}^T A\boldsymbol{x}/\|\boldsymbol{x}\|^2 \geq \lambda_n$ を示している。

   （A が半正値であることと、A の固有値はすべて非負であることが同値であることも同じ定理から得られる。）

   (i),(ii) ⇔ (iii) は教科書の Theorem 9.5.3 であるが証明が書かれていないので以下に示す。

   A_k で A の上から k 行、左から k 列でできる主小行列を表すものとする。$\boldsymbol{x} \in \mathrm{Span}(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k)$ とすると $\boldsymbol{x}_k$ は $k+1$ 行目以下は 0 であるような列ベクトルである。また $\boldsymbol{x}$ の最初の k 行目までとったものを $\boldsymbol{x}_k$ と表すとすると、$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{x}_k^T A_k\boldsymbol{x}_k$ である。つまり、上から k 行目までしか 0 でない成分が

ないベクトルに対応する $\boldsymbol{x}^T A\boldsymbol{x}$ は A_k だけで計算できることを意味している。したがって、A が正値ならば A_k も正値でなければならない。特に A_k の固有値はすべて 正となり、その積が行列式であったから $|A_k| > 0$（または $|A_k| \geq 0$）である。最後の部分は、$\det(T^{-1}AT) = \det(T^{-1})\det(A)\det(T) = \det(T^{-1})\det(T)\det(A) = \det(A)$ だから対角化た行列 $T^{-1}AT$ を考えれば（実対称行列はつねに直交対角化できたから）その対角線に並ぶものが固有値であることがわかり、それが正であることが分かる。これにより (i) $\Rightarrow$ (iii) が示された。

(iii) を仮定する。帰納法により A_{n-1} は正値であると仮定する。

$$A = \begin{bmatrix} A_{n-1} & \boldsymbol{b} \\ \boldsymbol{b}^T & c \end{bmatrix},\ P = \begin{bmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \boldsymbol{0}^T & 1 \end{bmatrix}\ \text{また}\ B = \begin{bmatrix} A_{n-1} & \boldsymbol{0} \\ \boldsymbol{0}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{bmatrix}$$

とすると $A = P^T BP$ であることが確かめられる。$\det(A) = \det(A_{n-1})(c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b})$。仮定より $\det(A) > 0$ かつ $\det(A_{n-1}) > 0$ だから $c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} > 0$ となる。$d = c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} > 0$ とおけば、A_{n-1} が正値なことから、B が正値となり、$A = P^T BP$ も正値となる。 ■

5. 3 から、(d) の行列に、2 の判定条件 (iii) を用いれば、正値でないことが分かる。また、$-A$ とすると、主小行列式の値は、4, -37, -1561 となるから、負値でもない。$|A| \neq 0$ だから半正値でも、半負値でもない。つまり、indefinite であることが分かった。$\det(-A_k) = (-1)^k \det(A_k)$ であることに注意。

9.5.17 *Complete the proof of Theorem 9.5.1 by showing that $\lambda_n \leq \boldsymbol{x}^T A\boldsymbol{x}$ if $\|\boldsymbol{x}\| = 1$ and $\lambda_n = \boldsymbol{x}^T A\boldsymbol{x}$ if $\boldsymbol{x}$ is an eigenvector of A corresponding to λ_n.*

Theorem 9.5.1 *Let A be a symmetric $n \times n$ matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. If $\boldsymbol{x}$ is constrained so that $\|\boldsymbol{x}\| = 1$, then*

(a) $\lambda_1 \geq \boldsymbol{x}^T A\boldsymbol{x} \geq \lambda_n$.

(b) $\boldsymbol{x}^T A\boldsymbol{x} = \lambda_n$ *if $\boldsymbol{x}$ is an eigenvector fo A corresponding to λ_n and $\boldsymbol{x}^T A\boldsymbol{x} = \lambda_1$ if $\boldsymbol{x}$ is an eigenvector of A corresonding to λ_1.*

Solution. Since A is a real symmetric matrix, it is diagonalizable by a real orthogonal matrix. Hence $V = \boldsymbol{R}^n$ has an orthonormal basis consisting of eivenvectors of A. Therefore $V = V(\lambda_1) + \cdots + V(\lambda_r)$, where $V(\lambda_i)$ is the eigenspace of A corresponding to λ_i, and for each $\boldsymbol{x} \in V$, there are $\boldsymbol{v}_i \in V(\lambda_i)$ such that $\boldsymbol{x} = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_r$, and $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ if $i \neq j$. Hence

$$\frac{\boldsymbol{x}^T A\boldsymbol{x}}{\|\boldsymbol{x}\|^2} = \frac{(\boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_r)^T(\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + \cdots + \lambda_r\boldsymbol{v}_r)}{\|\boldsymbol{x}\|^2} = \frac{\lambda_1\|\boldsymbol{v}_1\|^2 + \lambda_2\|\boldsymbol{v}_2\|^2 + \cdots + \lambda_r\|\boldsymbol{v}_r\|^2}{\|\boldsymbol{x}\|^2},$$

and

$$\lambda_1 = \frac{\lambda_1\|\boldsymbol{x}\|^2}{\|\boldsymbol{x}\|^2} \geq \frac{\lambda_1\|\boldsymbol{v}_1\|^2 + \lambda_2\|\boldsymbol{v}_2\|^2 + \cdots + \lambda_r\|\boldsymbol{v}_r\|^2}{\|\boldsymbol{x}\|^2} \geq \frac{\lambda_n\|\boldsymbol{x}\|^2}{\|\boldsymbol{x}\|^2} = \lambda_n,\ \text{as}$$

$$\|\boldsymbol{x}\|^2 = \|\boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_r\|^2 = \|\boldsymbol{v}_1\|^2 + \|\boldsymbol{v}_2\|^2 + \cdots + \|\boldsymbol{v}_r\|^2.$$

This proves (a) by setting $\|\boldsymbol{x}\| = 1$.

Moreover $\boldsymbol{x}^T A\boldsymbol{x} = \lambda_n$ if and only if $\boldsymbol{x} = \boldsymbol{v}_n$ and $\boldsymbol{x}^T A\boldsymbol{x} = \lambda_1$ if and only if $\boldsymbol{x} = \boldsymbol{v}_1$. ■

9.5.19 *Indicate whether each statement is true (T) or false (F). Justify your answer.*

(a) *If $\boldsymbol{x}$ is a vector in $\boldsymbol{R}^n$, then $\boldsymbol{x} \cdot \boldsymbol{x}$ is a quadratic form.*

(b) *If $\boldsymbol{x}^T A\boldsymbol{x}$ is a positive definite quadratic form, then so is $\boldsymbol{x}^T A^{-1}\boldsymbol{x}$.*

(c) *If A is a matrix with positive eigenvalues, then $\boldsymbol{x}^T A\boldsymbol{x}$ is a positive quadratic form.*

(d) *If A is a symmetrix 2×2 matrix with positive entries and a positive determinant, then A is positive definite.*

(e) *If $\boldsymbol{x}^T A\boldsymbol{x}$ is a quadratic form with no cross-product terms, then A is a diagonal matrix.*

(f) *If $\boldsymbol{x}^T A\boldsymbol{x}$ is a positive definite quadratic form in x and y, and if $c \neq 0$, then the graph of the equation $\boldsymbol{x}^T A\boldsymbol{x} = c$ is an ellipse.*

Solution.

(a) (T). The quadratic form can be written $x_1^2+x_2^2+\cdots+x_n^2$ and the corresponding matrix representing it is the identity matrix I_n.

(b) (T). Since $(A^{-1})^T = (A^T)^{-1} = A^{-1}$, A^{-1} is a symmetric matrix. λ_i is an eigenvalue of A if and only if λ_i^{-1} is an eigenvalue of A^{-1} as $A\boldsymbol{x} = \lambda\boldsymbol{x}$ implies $\lambda^{-1}\boldsymbol{x} = A^{-1}\boldsymbol{x}$. Note that since A is positive definite, every eigenvalue $\lambda_i > 0$ and hence $\lambda_i^{-1} > 0$.

(c) (F). A has to be a real symmetric matrix.

(d) (T). Since A is a symmetric 2×2 matrix, the eigenvalues are real. Let λ and μ be eigenvalues. Then $\lambda+\mu = \mathrm{tr}(A)$ and $\lambda\cdot\mu = \det(A)$ and these two values are positive by our assumption. Hence $\lambda > 0$ and $\mu > 0$ and A is positive definite. This is also a direct consequence of Theorem 9.5.3. If A is a symmetrix 2×2 matrix with positive entries and a positive determinant, then A is positive definite.

(e) (T). This is clear from the correspondence between A and the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$.

(f) (F). If $c < 0$, then there is no $\boldsymbol{x}$ satisfies the condition. If $c > 0$, there is an orthogonal matrix T such that $T^T AT$ is a diagonal matrix D with positive diagonal entries. Hence by changing the basie if $\boldsymbol{x}' = T^T\boldsymbol{x}$, then $\boldsymbol{x}'^T D\boldsymbol{x}' = \boldsymbol{x}^T TDT^T\boldsymbol{x} = \boldsymbol{x}^T T\boldsymbol{x}$. If $D = \mathrm{diag}(\alpha,\beta)$ and $\boldsymbol{x}' = [x', y']^T$, then $\boldsymbol{x}'^T D\boldsymbol{x}' = \alpha x'^2 + \beta y'^2$ and this is equal to c. Therefore the graph of $\alpha x'^2 + \beta y'^2 = c$ represents an ellipse as $\alpha > 0$ and $\beta > 0$.

まとめ

1. T を正則行列としたとき、A と $T^{-1}AT$ を相似 (similar) な行列と言います。A の性質を見るために、より簡単な形の行列で A と相似なもの $T^{-1}AT$ を考え、その性質から A の性質を導くことがこの授業の主題でした。そのためには、相似変換をしたときに、変わらない性質が重要でした。そこで、そのような性質をまとめてみましょう。(8.5 参照)

2. $\det(T^{-1}AT) = \det(A)$, $\mathrm{tr}(T^{-1}AT) = \mathrm{tr}(A)$。前者は行列式の積が行列の積の行列式であることから、後者は、$\mathrm{tr}(AB) = \mathrm{tr}(BA)$ から得られます。

3. $\mathrm{char}_{T^{-1}AT}(\lambda) = \mathrm{char}_A(\lambda)$。（この性質は基本的に、$\det(T^{-1}XT) = \det(X)$ から導かれます。）この事から、A の固有値と、$T^{-1}AT$ の固有値は重複度（algebraic multiplicity）を含めて同じであることが分かります。特に、$T^{-1}AT$ が上半三角行列（または下半三角行列）であれば、対角線上に固有値が並びますが、それは、A の固有値であることも分かります。

4. A を $L: V \to V$ $(\boldsymbol{x} \mapsto A\boldsymbol{x})$ なる線形変換 (linear transformation) の標準基底 $B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ に関する行列表示だと考えると、$T^{-1}AT$ は別の基底 $B' = \{T\boldsymbol{e}_1, T\boldsymbol{e}_2, \ldots, T\boldsymbol{e}_n\}$ に関する行列表現でした。$(L_{B'} = [[L(T\boldsymbol{e}_1)]_{B'}, [L(T\boldsymbol{e}_2)]_{B'}, \ldots, [L(T\boldsymbol{e}_n)]_{B'}] = [[AT\boldsymbol{e}_1]_{B'}, [AT\boldsymbol{e}_2]_{B'}, \ldots, [AT\boldsymbol{e}_n]_{B'}] = [[T^{-1}AT\boldsymbol{e}_1]_B, [T^{-1}AT\boldsymbol{e}_2]_B, \ldots, [T^{-1}AT\boldsymbol{e}_n]_B] = T^{-1}AT)$

5. このことから、A の固有値 λ に関する固有空間の次元と、$T^{-1}AT$ の同じ固有値 λ に関する固有空間の次元は同じになります。実際 $V(\lambda) = \{\boldsymbol{v} \mid A\boldsymbol{v} = \lambda\boldsymbol{v}\}$ とすると、

$$T^{-1}V(\lambda) = \{T^{-1}\boldsymbol{v} \mid A\boldsymbol{v} = \lambda\boldsymbol{v}\} = \{\boldsymbol{w} \mid T^{-1}AT\boldsymbol{w} = \lambda\boldsymbol{w}\}$$

となり、T^{-1} が可逆なことから、$\dim V(\lambda) = \dim\{\boldsymbol{w} \mid T^{-1}AT\boldsymbol{w} = \lambda\boldsymbol{w}\}$ が分かります。

6. $\boldsymbol{x} \in V$ は $T\boldsymbol{y}$ $(\boldsymbol{y} \in V)$ とも表せるので、$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{y}^T T^T AT\boldsymbol{y}$ となり、T が直交行列のときは、A が正値であることと、$T^{-1}AT$ が正値であること、などが同値になります。このことは、固有値による正値の判定条件からも明らかです。