To obtain an eigenvector $x^2 \neq 0_2$ for $\lambda_2 = 0$, define:

$$B^2 \equiv A - \lambda_2 I_2 = \begin{bmatrix} 1, & 1 \\ 1, & 1 \end{bmatrix} - 0 \begin{bmatrix} 1, & 0 \\ 0, & 1 \end{bmatrix} = \begin{bmatrix} 1, & 1 \\ 1, & 1 \end{bmatrix}$$

To transform B^2 into an upper triangular matrix, subtract the first row of B^2 from the second row of B^2 and we obtain U^2:

$$U^2 \equiv \begin{bmatrix} 1, & 1 \\ 0, & 0 \end{bmatrix}$$

The first zero element of U^2 is $u_{22}^2 = 0$. Hence set $x_2^2 = 1$ and solve

$$U^2 x^2 = \begin{bmatrix} 1, & 1 \\ 0, & 0 \end{bmatrix} \begin{bmatrix} x_1^2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

for $x_1^2 = -1$. Hence $x^2 = [-1, 1]^T$ is an eigenvector of A that corresponds to the eigenvalue $\lambda_2 = 0$.

Note that the two eigenvectors x^1 and x^2 that were constructed in Examples 2 and 3 above had the property.

(139) $x^{1T}x^2 = 0$.

This property turns out to be a general property of eigenvectors of a symmetric A that correspond to distinct eigenvalues as we shall see later.

Problems:

17. Calculate the 2 eigenvalues λ_1 and λ_2 of $A \equiv \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ and calculate eigenvectors x^1 and x^2 that correspond to λ_1 and λ_2.

18. Calculate the 3 eigenvalues and eigenvectors of $A \equiv \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$.

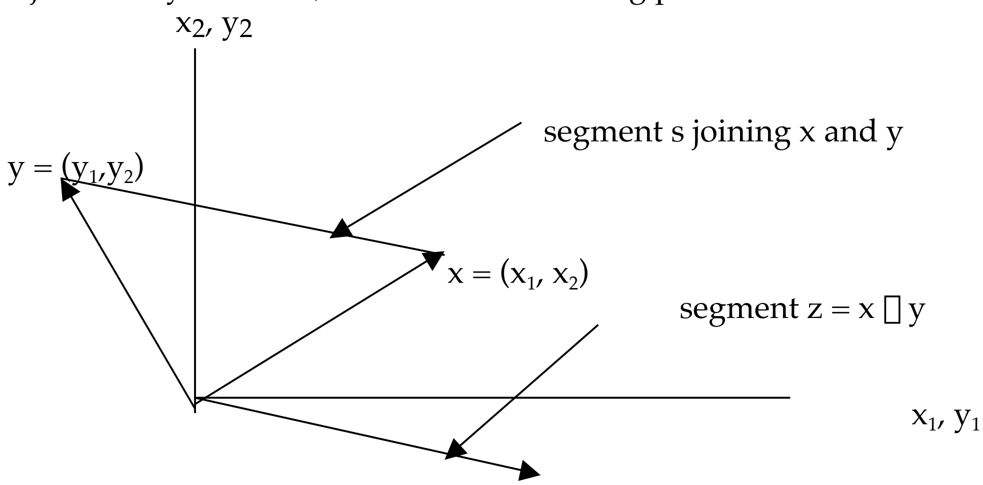
Hint: Taking into account the block diagonal structure of A, you may find the results of Problem 17 useful.

(140) **Definition:** Two N dimensional vectors x and y are *orthogonal* or *perpendicular* iff $x^Ty = 0$.

To see why $x^Ty = 0$ implies x and y are perpendicular, let x and y be two N dimensional vectors of unit length; i.e.,

(141) $x^Tx = \sum_{i=1}^{N} x_i^2 = 1$ and $y^Ty = \sum_{i=1}^{N} y_i^2 = 1.$

Now look at the vector $z \equiv x - y$ which has the same length as the line segment s which joins x to y. If N = 2, we have the following picture.



If x and y are perpendicular and of unit length, then by Pythagoras' Theorem in geometry, the length of the segment s (which is equal to the length of z) is $\sqrt{2}$; i.e., we have

(142) $z^Tz = x^Tx + y^Ty = 1 + 1 = 2.$

Now substitute $z = x - y$ into (142) and simplify:

$$\begin{aligned} 2 &= (x - y)^T (x - y) \\ &= x^Tx - x^Ty - y^Tx + y^Ty \\ &= 2 - x^Ty - y^Tx && \text{using (141)} \\ \text{(143)} \quad &= 2 - 2x^Ty && \text{since } x^Ty = y^Tx. \end{aligned}$$

Equation (143) implies that $x^Ty = 0$. Thus if x and y are of unit length and are perpendicular, then $x^Ty = 0$. This argument can be extended to the case where $x \neq 0_N$ and $y \neq 0_N$. (If x and y are perpendicular, then so are $x/(x^Tx)^{1/2} \equiv a$ and $y/(y^Ty)^{1/2} \equiv b$. Now a and b are perpendicular and of unit length and so we can apply the above argument to conclude that $a^Tb = 0$. But this implies $x^Ty = 0$ as well).

(144) **Theorem:** Suppose that the N eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ of the N by N symmetric matrix A are all *distinct*. Let $x^1, x^2, \ldots, x^N$ be eigenvectors corresponding to these distinct eigenvalues. Then $x^{iT}x^j = 0$ for all $i \neq j$; i.e., the eigenvectors $x^1, x^2, \ldots, x^N$ are mutually *orthogonal*.

Proof: Let $i \neq j$. By the definition of $x^i \neq 0_N$ and λ_i being an eigenvalue and eigenvector of A, we have:

(145) $Ax^i = \lambda_i x^i$

and by the definition of λ_j and $x^j \neq 0_N$ being an eigenvalue and eigenvector of A, we have:

(146) $Ax^j = \lambda_j x^j$.

Premultiply both sides of (145) by x^{jT} and obtain:

(147) $x^{jT}Ax^i = \lambda_i x^{jT}x^i$.

Now take transposes of both sides of (147) and get:

(148) $\lambda_i x^{iT}x^j = x^{iT}A^Tx^j = x^{iT}Ax^j$

where the second equality in (148) follows from $A = A^T$. Premultiply both sides of (146) by x^{iT} and obtain:

(149) $\lambda_j x^{iT} x^j = x^{iT}Ax^j$.

Since the right hand sides of (148) and (149) are equal, so are the left hand sides, so we obtain:

(150) $\lambda_i x^{iT}x^j = \lambda_j x^{iT}x^j$ or

(151) $(\lambda_i - \lambda_j)x^{iT} x^j = 0$.

Since $\lambda_i \neq \lambda_j$ by assumption, (151) implies that $x^{iT}x^j = 0$; i.e., x^i is orthogonal to x^j.

Q.E.D.

10. The Diagonalization of a Symmetric Matrix by an Orthonormal Transformation

Suppose that A is an N by N symmetric matrix with *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ with corresponding nonzero eigenvectors $x^1, x^2, \ldots, x^N$. (We will deal with the case where the eigenvalues are not necessarily distinct later in this section). We *normalize* the eigenvectors $x^i \neq 0_N$ so that they are of unit length; i.e., for i = 1, 2, . . ., N, define the *normalized eigenvector* u^i by

(152) $u^i = x^i/(x^{iT}x^i)^{1/2}$; $\quad i = 1, 2, \ldots, N$.

It can be seen that since $Ax^i = \lambda_i x^i$ for i = 1, . . ., N, we also have:

(153) $Au^i = \lambda_i u^i$; $\qquad$ i = 1, 2, . . ., N,

and the u^i also satisfy:

(154) $$\begin{aligned} u^{iT}u^i &= [x^i/(x^{iT}x^i)^{1/2}]^T [x^i/(x^{iT}x^i)^{1/2}] \\ &= x^{iT}x^i/(x^{iT}x^i)^{1/2+1/2} \\ &= 1 \qquad \text{for } i = 1, \ldots, N. \end{aligned}$$

From Theorem (144), we have $x^{iT}x^j = 0$ if $i \neq j$. This implies that the u^i have the same orthogonality properties; i.e., we have:

(155) $u^{iT}u^j = 0$ $\quad$ if $\quad i \neq j$.

Define the N by N matrix U of the normalized eigenvectors of A by:

(156) $U \equiv [u^1, u^2, \ldots, u^N]$.

Using (154) and (155), it can be seen that

(157) $U^TU = I_N$.

But (157) tells us that U^T is a left inverse (and hence is an inverse) for U; i.e.,

(158) $U^{-1} = U^T$.

Taking determinants of both sides of (157) yields:

(158) $1 = |I_N| = |U^TU| = |U^T|\,|U| = |U|^2$, or

(159) $|U| = +1$ or -1.

A square matrix U that satisfies (157) is called an *orthonormal matrix*, and (159) shows us that its determinant equals +1 or -1.

Return to the N eigenvalue, normalized eigenvector equations (153). Using definition (156), it can be seen that the N equations (153) can be rewritten as the following matrix equation:

$$\begin{aligned} AU &= [\lambda_1 u^1, \lambda_2 u^2, \ldots, \lambda_N u^N] \\ &= [u^1, u^2, \ldots, u^N] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix} \end{aligned}$$

(160) $\quad = U\Lambda$

where Λ is a diagonal matrix with the eigenvalues of A on the main diagonal of Λ. Now premultiply both sides of (160) by U^T and get:

(161) $\quad U^TAU = U^TU\Lambda = \Lambda \qquad$ using (157)

We can now use the eigenvector-eigenvalue method for diagonalizing A, the matrix equation (161) above, in exactly the same way that we used the Gauss-Lagrange diagonalization method (59) above in order to determine the definiteness properties of A. For example, since $|U| \neq 0$, it can be seen that the conditions for positive definiteness of A,

(162) $\quad x^TAx > 0 \qquad$ for all $x \neq 0_N$

are equivalent to

$$\begin{aligned} 0 &< x^TAx, && x \neq 0_N \\ &= y^TU^TAUy, && \text{letting } x = Uy \text{ so } y \neq 0_N \\ &= y^T \Lambda y, && \text{using (161)} \\ &= \textstyle\sum_{i=1}^N \lambda_i y_i^2. && (163) \end{aligned}$$

Thus A is positive (negative) definite iff all of the eigenvalues of A, $\lambda_1, \lambda_2, \ldots, \lambda_N$, are positive (negative). Of course A is positive (negative) semidefinite iff all of the eigenvalues of A are nonnegative (nonpositive).

However, we established the above results under the assumption that all of the eigenvalues of A were *distinct*. In the next section, we shall relax this assumption that the eigenvalues of A are distinct, but we will still get the same results as in the above paragraph.

11. The Diagonalization of a Symmetric Matrix in the General Case

In order to derive the matrix equation (161) in the previous section without the assumption that the eigenvalues of the symmetric N by N matrix are distinct, we need two preliminary results.

(164) *The Gram-Schmidt Orthogonalization Procedure*: Let the N dimensional vectors $x^1, x^2, \ldots, x^K$ be linearly independent (so that $K \leq N$). Then for $k = 1, 2, \ldots, K$; each vector x^K can be expressed as a linear combination of k *orthonormal* vectors $y^1, y^2, \ldots, y^K$; i.e., for $k = 1, 2, \ldots, K$:

(165) $\quad x^k = \sum_{j=1}^k a_{kj}y^j \qquad$ where

(166) $y^{iT}y^j = \begin{cases} 1 & \text{if} \quad i = j \quad \text{and} \\ 0 & \text{if} \quad i \neq j. \end{cases}$

The a_{kj} in equations (165) are scalars.

Proof: For k = 1, take $y^1 \equiv x^1/(x^{1T}x^1)^{1/2}$. Note that $x^{1T}x^1 > 0$ since if $x^1 = 0_N$, then the x^k would not be linearly independent. Thus we have $x^1 = a_{11}y^1$ where $a_{11} \equiv (x^{1T}x^1)^{1/2}$.

For k = 2, define the vector z^2 as follows:

(167) $z^2 \equiv x^2 - (x^{2T}y^1)y^1.$

Suppose $z^2 = 0_N$. Then $x^2 = (x^{2T}y^1)y^1 = x^1(x^{2T}y^1)/(x^{1T}x^1)^{1/2}$

which would imply that x^2 is a multiple of x^1 and hence x^1 and x^2 would be linearly dependent. Thus our *supposition* is false and $z^2 \neq 0_N$.

We show that z^2 is orthogonal to y^1:

(168)
$$\begin{aligned} y^{1T}z^2 &= y^{1T}[x^2 - (x^{2T}y^1)y^1] \\ &= y^{1T}x^2 - (x^{2T}y^1)y^{1T}y^1 \\ &= y^{1T}x^2 - (x^{2T}y^1) && \text{using } y^{1T}y^1 = 1 \\ &= 0 && \text{since } (x^{2T}y^1)^T = y^{1T}x^2. \end{aligned}$$

Now define y^2 as the following normalization of z^2:

(169) $y^2 \equiv z^2/(z^{2T}z^2)^{1/2}$

and so we have using (168) and (169)):

(170) $y^{1T}y^2 = 0,\ y^{2T}y^2 = 1$ and $y^{1T}y^1 = 1$.

If we substitute (169) into (167) and rearrange terms, we have:

(171)
$$\begin{aligned} x^2 &= (x^{2T}y^1)y^1 = (z^{2T}z^2)^{1/2}y^2 \\ &\equiv a_{21}y^1 + a_{11}\,y^2 \end{aligned}$$

which is (165) for k = 2.

For a general k, having defined $y^1, y^2, \ldots, y^{k-1}$, define z^k as follows:

(172) $z^k \equiv x^k - \sum_{j=1}^{k-1} (x^{kT}y^j)y^j.$

If $z^k = 0_N$, then we can show that $x^1, x^2, \ldots, x^k$ are linearly dependent which contradicts our assumption. Thus $z^k \neq 0_N$. It is also straightforward to show that for $l = 1, 2, \ldots, k-1$:

$$
\begin{aligned}
(173)\quad y^{lT}z^k &= y^{lT}[x^k - \textstyle\sum_{j=1}^{k-1}(x^{kT}y^j)y^j] \\
&= y^{lT}x^k - \textstyle\sum_{j=1}^{k-1}(x^{kT}y^j)y^{lT}y^j \\
&= y^{lT}x^k - (x^{kT}y^l)y^{lT}y^l && \text{since } y^{lT}y^j = 0 \quad \text{if } l \neq j \\
&= y^{lT}x^k - (x^{kT}y^k) && \text{since } y^{lT}y^l = 1 \\
&= 0.
\end{aligned}
$$

Thus z^k is orthogonal to $y^l, y^2, \ldots, y^{k-1}$. Now define y^k as the following normalization of z^k:

$$(174)\quad y^k \equiv z^k/(z^{kT}z^k)^{1/2}.$$

Now substitute (174) into (172) and we obtain:

$$
\begin{aligned}
(175)\quad x^k &= \textstyle\sum_{j=1}^{k-1}(x^{kT}y^j)y^j + (z^{kT}z^k)^{1/2}y^k \\
&\equiv \textstyle\sum_{j=1}^{k-1} a_{kj}y^j + a_{kk}\, y^k.
\end{aligned}
$$

Q.E.D.

The Gram-Schmidt orthogonalization procedure is very useful in econometrics.

We need another preliminary result about orthonormal N by N matrices U (recall this means $U^TU = I_N$).

(176) **Lemma:** Let U^1 and U^2 be two N by N orthonormal matrices. Then the product matrix $U \equiv U^1U^2$ is also orthonormal.

Proof: $U^TU = (U^1U^2)^T(U^1U^2) = U^{2T}U^{1T}U^1U^2 = U^{2T}I_NU^2 = I_N$.

Q.E.D.

Now we are ready to derive the diagonal decomposition of the N by N matrix A, $U^TAU = \Lambda$ where $U^TU = I_N$, assuming only that A is symmetric; i.e., we want to drop our hypothesis that we made in the previous section that the eigenvalues of A were distinct.

Let A be N by N and symmetric and let λ_1 be any eigenvalue of A with corresponding normalized eigenvector y^1; i.e., we have

$$(177)\quad Ay^1 = \lambda_1 y^1 \quad \text{with} \quad y^{1T}y^1 = 1.$$