Lecture 20

Orthogonal matrices and orthonormal lists

Definition 23.1: Let A be an $n \times n$ matrix.
We say that A is an *orthogonal matrix* if $A^T A = I_n$,
or equivalently A is invertible and $A^{-1} = A^T$.

Definition 23.2: Let $v_1, \dots, v_r$ be a list of r vectors in $\mathbb{R}^n$.
We say that this list is *orthonormal* if $v_i.v_i = 1$ for all i,
and $v_i.v_j = 0$ whenever i and j are different.

Proposition 23.4: Any orthonormal list of length n in $\mathbb{R}^n$ is a basis.

Proof: Let $v_1, \dots, v_n$ be an orthonormal list of length n.
Suppose we have a linear relation $\sum_{i=1}^n \lambda_i v_i = 0$.
We can take the dot product of both sides with v_p to get $\sum_{i=1}^n \lambda_i (v_i.v_p) = 0$.
Most of the terms $v_i.v_p$ are zero, because $v_i.v_j = 0$ whenever $i \neq j$.
After dropping the terms where $i \neq p$, we are left with $\lambda_p (v_p.v_p) = 0$.
Here $v_p.v_p = 1$ (by the definition of orthonormality) so $\lambda_p = 0$.
This works for all p, so our linear relation is the trivial one.
This proves that the list $v_1, \dots, v_n$ is linearly independent.
A linearly independent list of n vectors in $\mathbb{R}^n$ is automatically a basis by Proposition 10.12. □

Orthogonal matrices and orthonormal lists

Proposition 23.5: Let A be an $n \times n$ matrix. Then A is an orthogonal matrix if and only if the columns of A form an orthonormal list.

Proof.
By definition, A is orthogonal if and only if A^T is an inverse for A, or in other words $A^T A = I_n$. Let the columns of A be $v_1, \dots, v_n$. Then

$$A^T A = \begin{bmatrix} v_1^T \\ \hline \vdots \\ \hline v_n^T \end{bmatrix} \begin{bmatrix} v_1 & \vert & \cdots & \vert & v_n \end{bmatrix} = \begin{bmatrix} v_1.v_1 & \cdots & v_1.v_n \\ \vdots & \ddots & \vdots \\ v_n.v_1 & \cdots & v_n.v_n \end{bmatrix}$$

In other words, the entry in the (i,j) position in $A^T A$ is just the dot product $v_i.v_j$. For $A^T A$ to be the identity we need the diagonal entries $v_i.v_i$ to be one, and the off-diagonal entries $v_i.v_j$ (with $i \neq j$) to be zero. This means precisely that the list $v_1, \dots, v_n$ is orthonormal. □

Symmetric matrices

Definition 23.6: Let A be an $n \times n$ matrix, with entries a_{ij}.
We say that A is *symmetric* if $A^T = A$,
or equivalently $a_{ij} = a_{ji}$ for all i and j.

Example: A 4×4 matrix is symmetric if and only if it has the form

$$\begin{bmatrix} a & b & c & d \\ b & e & f & g \\ c & f & h & i \\ d & g & i & j \end{bmatrix}.$$

Example: The matrices A and B are symmetric, but C and D are not.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 111 & 11 & 1 \\ 11 & 111 & 11 \\ 1 & 11 & 111 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 10 & 1000 \\ 1 & 10 & 1000 \\ 1 & 10 & 1000 \end{bmatrix}$$

Dot products and symmetric matrices

Lemma 23.9: Let A be an $n \times n$ matrix, and let u and v be vectors in $\mathbb{R}^n$. Then $u.(Av) = (A^T u).v$. Thus, if A is symmetric then $u.(Av) = (Au).v$.

Proof.
Put $p = A^T u$ and $q = Av$, so the claim is that $u.q = p.v$.
By the definition of matrix multiplication, we have $q_i = \sum_j A_{ij} v_j$,
so $u.q = \sum_i u_i q_i = \sum_{i,j} u_i A_{ij} v_j$.
Similarly, we have $p_j = \sum_i (A^T)_{ji} u_i$, but $(A^T)_{ji} = A_{ij}$ so $p_j = \sum_i u_i A_{ij}$.
It follows that $p.v = \sum_j p_j v_j = \sum_{i,j} u_i A_{ij} v_j$,
which is the same as $u.q$, as claimed.

Alternatively: for $x, y \in \mathbb{R}^n$ the dot product $x.y$ is the matrix product $x^T y$.
Thus $(Au).v = (Au)^T v$, but $(Au)^T = u^T A^T$ (by Proposition 3.4)
so $(Au).v = u^T (A^T v) = u.(A^T v)$. □

Eigenvalues of symmetric matrices

Proposition 23.10: Let A be an $n \times n$ symmetric matrix (with real entries).

(a) All eigenvalues of A are real numbers.

(b) If u and v are (real) eigenvectors for A with distinct eigenvalues, then u and v are orthogonal.

Proof of (a): Let $\lambda = \alpha + i\beta$ be a complex eigenvalue of A ($\alpha, \beta \in \mathbb{R}$).
We must show that $\beta = 0$, so that λ is actually a real number.
As λ is an eigenvalue, there is a nonzero vector u with $Au = \lambda u$.
Let $v, w \in \mathbb{R}^n$ be the real and imaginary parts of u, so $u = v + iw$.

$$Av + iAw = A(v + iw) = Au = \lambda u = (\alpha + i\beta)(v + iw) = (\alpha v - \beta w) + i(\beta v + \alpha w).$$

As the entries in A are real, we see that the vectors Av and Aw are real.
Compare real and imaginary parts to get

$$Av = \alpha v - \beta w \qquad Aw = \beta v + \alpha w$$

$$(Av).w = \alpha v.w - \beta w.w \qquad v.(Aw) = \beta v.v + \alpha v.w.$$

However, A is symmetric, so $(Av).w = v.(Aw)$ by Lemma 23.9.
Rearrange to get $\beta(v.v + w.w) = 0$ or $\beta(\|v\|^2 + \|w\|^2) = 0$.
By assumption $u \neq 0$ so ($v \neq 0$ or $w \neq 0$) so $\|v\|^2 + \|w\|^2 > 0$.
Divide by this to get $\beta = 0$ and $\lambda = \alpha \in \mathbb{R}$ as claimed.

Eigenvalues of symmetric matrices

Proposition 23.10: Let A be an $n \times n$ symmetric matrix (with real entries).

(a) All eigenvalues of A are real numbers.

(b) If u and v are (real) eigenvectors for A with distinct eigenvalues, then u and v are orthogonal.

Proof of (b): Suppose that u and v are eigenvectors of A with distinct eigenvalues, say λ and μ. This means that

$$Au = \lambda u \qquad Av = \mu v \qquad \lambda \neq \mu.$$

As A is symmetric we have $(Au).v = u.(Av)$.
As $Au = \lambda u$ and $Av = \mu v$ this becomes $\lambda\, u.v = \mu\, u.v$.
Rearrange to get $(\lambda - \mu)u.v = 0$.
As $\lambda \neq \mu$ we can divide by $\lambda - \mu$ to get $u.v = 0$,
which means that u and v are orthogonal. □

Alternative proof for 2×2 matrices

A 2×2 symmetric matrix has the form

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \qquad \text{so} \qquad A - tI_2 = \begin{bmatrix} a - t & b \\ b & d - t \end{bmatrix}$$

so

$$\chi_A(t) = (a - t)(d - t) - b^2 = t^2 - (a + d)t + (ad - b^2).$$

The eigenvalues are

$$\lambda = \frac{a + d \pm \sqrt{(a + d)^2 - 4(ad - b^2)}}{2}.$$

The expression under the square root is

$$\begin{aligned}(a + d)^2 - 4(ad - b^2) &= a^2 + 2ad + d^2 - 4ad + 4b^2 \\ &= a^2 - 2ad + d^2 + 4b^2 \\ &= (a - d)^2 + (2b)^2.\end{aligned}$$

This is the sum of two squares, so it is nonnegative, so the square root is real, so the two eigenvalues are both real.

Orthonormal basis of eigenvectors

Proposition 23.12: Let A be an $n \times n$ symmetric matrix.
Then there is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors for A.

Partial proof.
We will show that the Theorem holds whenever A has n distinct eigenvalues.
In fact it is true even without that assumption, but the proof is harder.

Let the eigenvalues of A be $\lambda_1, \dots, \lambda_n$ (so $\lambda_i \in \mathbb{R}$).
For each i we choose a (real) eigenvector u_i of eigenvalue λ_i.
As u_i is an eigenvector we have $u_i \neq 0$ and so $u_i.u_i > 0$
so we can define $v_i = u_i/\sqrt{u_i.u_i}$. This is just a real number times u_i, so it is again an eigenvector of eigenvalue λ_i.
It satisfies $v_i.v_i = \dfrac{u_i.u_i}{\sqrt{u_i.u_i}\,\sqrt{u_i.u_i}} = 1$ (so it is a unit vector).
Proposition 23.10(b): eigenvectors of a symmetric matrix with distinct eigenvalues are orthogonal. Thus $v_i.v_j = 0$ for $i \neq j$.
This shows that the sequence $v_1, \dots, v_n$ is orthonormal.
Proposition 23.4: any orthonormal list of length n in $\mathbb{R}^n$ is a basis.
Proposition 13.22: any n eigenvectors in $\mathbb{R}^n$ with distinct eigenvalues form a basis.
Either of these results implies that $v_1, \dots, v_n$ is a basis. □

Our special case is the usual case

Let A be an $n \times n$ symmetric matrix again.
The characteristic polynomial $\chi_A(t)$ has degree n,
so by well-known properties of polynomials it can be factored as

$$\chi_A(t) = \prod_{i=1}^{n} (\lambda_i - t)$$

for some complex numbers $\lambda_1, \dots, \lambda_n$.

By Proposition 23.10(a) these eigenvalues λ_i are in fact all real.

Some of them might be the same, but that would be a concidence which could only happen if the matrix A was very simple or had some kind of hidden symmetry.

Thus, our proof of Proposition 23.12 covers almost all cases
(but some of the cases that are not covered are the most interesting ones).

Orthonormal eigenvector example

Consider the symmetric matrix $A = \begin{bmatrix} 1&1&1&1&1 \\ 1&1&1&1&1 \\ 1&1&1&1&1 \\ 1&1&1&1&1 \\ 1&1&1&1&1 \end{bmatrix}$
(which appeared on one of the problem sheets) and the vectors

$$u_1 = \begin{bmatrix} 1\\-1\\0\\0\\0 \end{bmatrix} \quad u_2 = \begin{bmatrix} 1\\0\\-1\\0\\0 \end{bmatrix} \quad u_3 = \begin{bmatrix} 1\\0\\0\\-1\\0 \end{bmatrix} \quad u_4 = \begin{bmatrix} 1\\0\\0\\0\\-1 \end{bmatrix} \quad u_5 = \begin{bmatrix} 1\\1\\1\\1\\1 \end{bmatrix}.$$

These satisfy $Au_1 = Au_2 = Au_3 = Au_4 = 0$ and $Au_5 = 5u_5$,
so they are eigenvectors of eigenvalues $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ and $\lambda_5 = 5$.
Because λ_5 is different from $\lambda_1, \dots, \lambda_4$, Proposition 23.10(b) tells us that u_5 must be orthogonal to $u_1, \dots, u_4$, and indeed it is easy to see directly that $u_1.u_5 = \cdots = u_4.u_5 = 0$. However, the eigenvectors $u_1, \dots, u_4$ all share the same eigenvalue so there is no reason for them to be orthogonal and in fact they are not.

Orthonormal eigenvector example

$$u_1 = \begin{bmatrix} 1\\-1\\0\\0\\0 \end{bmatrix} \quad u_2 = \begin{bmatrix} 1\\0\\-1\\0\\0 \end{bmatrix} \quad u_3 = \begin{bmatrix} 1\\0\\0\\-1\\0 \end{bmatrix} \quad u_4 = \begin{bmatrix} 1\\0\\0\\0\\-1 \end{bmatrix} \quad u_5 = \begin{bmatrix} 1\\1\\1\\1\\1 \end{bmatrix}.$$

These satisfy $Au_1 = Au_2 = Au_3 = Au_4 = 0$ and $Au_5 = 5u_5$.

The vectors $u_1, \dots, u_4$ are not orthogonal:

$$u_1.u_2 = u_1.u_3 = u_1.u_4 = u_2.u_3 = u_2.u_4 = u_3.u_4 = 1.$$

However, it is possible to choose a different basis of eigenvectors where all the eigenvectors are orthogonal to each other. One such choice is as follows:

$$v_1 = \begin{bmatrix} 1\\-1\\0\\0\\0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 1\\1\\-2\\0\\0 \end{bmatrix} \quad v_3 = \begin{bmatrix} 1\\1\\1\\-3\\0 \end{bmatrix} \quad v_4 = \begin{bmatrix} 1\\1\\1\\1\\-4 \end{bmatrix} \quad v_5 = \begin{bmatrix} 1\\1\\1\\1\\1 \end{bmatrix}$$

It is easy to check directly that $Av_1 = Av_2 = Av_3 = Av_4 = 0$ $\quad Av_5 = 5v_5$

$$v_1.v_2 = v_1.v_3 = v_1.v_4 = v_1.v_5 = v_2.v_3 = v_2.v_4 = v_2.v_5 = v_3.v_4 = v_3.v_5 = v_4.v_5 = 0,$$

so the v_i are eigenvectors and are orthogonal to each other.

Orthonormal eigenvector example

$$v_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \\ 0 \end{bmatrix} \qquad v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -3 \\ 0 \end{bmatrix} \qquad v_4 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -4 \end{bmatrix} \qquad v_5 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

These satisfy $Av_1 = Av_2 = Av_3 = Av_4 = 0$ and $Av_5 = 5v_5$, and they are orthogonal to each other.

However, the list $v_1, \dots, v_5$ is not orthonormal, because

$$v_1.v_1 = 2 \qquad v_2.v_2 = 6 \qquad v_3.v_3 = 12 \qquad v_4.v_4 = 20 \qquad v_5.v_5 = 5.$$

This is easily fixed: if we put

$$w_1 = \frac{v_1}{\sqrt{2}} \qquad w_2 = \frac{v_2}{\sqrt{6}} \qquad w_3 = \frac{v_3}{\sqrt{12}} \qquad w_4 = \frac{v_4}{\sqrt{20}} \qquad w_5 = \frac{v_5}{\sqrt{5}}$$

then $w_1, \dots, w_5$ is an orthonormal basis for $\mathbb{R}^5$ consisting of eigenvectors for A.

Orthogonal diagonalisation of symmetric matrices

Corollary 23.15: Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal matrix U and a diagonal matrix D such that $A = UDU^T = UDU^{-1}$.

Proof.
Choose an orthonormal basis of eigenvectors $u_1, \dots, u_n$, and let λ_i be the eigenvalue of u_i.

Put $U = [u_1 | \cdots | u_n]$ and $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Proposition 14.4 tells us that $U^{-1}AU = D$ and so $A = UDU^{-1}$.

Proposition 23.5 tells us that U is an orthogonal matrix, so $U^{-1} = U^T$. □

Example of orthogonal diagonalisation

Let A be the 5×5 matrix in which every entry is one, as in Example 23.14. Following the prescription in the above proof, we put

$$U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ 0 & -2/\sqrt{6} & 1/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ 0 & 0 & -3/\sqrt{12} & 1/\sqrt{20} & 1/\sqrt{5} \\ 0 & 0 & 0 & -4/\sqrt{20} & 1/\sqrt{5} \end{bmatrix} \qquad D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

The general theory tells us that $A = UDU^T$. We can check this directly:

$$UD = \begin{bmatrix} * & * & * & * & 1/\sqrt{5} \\ * & * & * & * & 1/\sqrt{5} \\ * & * & * & * & 1/\sqrt{5} \\ * & * & * & * & 1/\sqrt{5} \\ * & * & * & * & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \end{bmatrix}$$

$$UDU^T = \begin{bmatrix} 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \\ 0 & 0 & 0 & 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ 1/\sqrt{5} & 1/\sqrt{5} & 1/\sqrt{5} & 1/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = A.$$

Example of orthogonal diagonalisation

Write $\rho = \sqrt{3}$ for brevity (so $\rho^2 = 3$), and consider the symmetric matrix

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

The characteristic polynomial is

$$\begin{aligned} \chi_A(t) &= \det \begin{bmatrix} -t & 1 & \rho \\ 1 & -t & -\rho \\ \rho & -\rho & -t \end{bmatrix} \\ &= -t \det \begin{bmatrix} -t & -\rho \\ -\rho & -t \end{bmatrix} - \det \begin{bmatrix} 1 & -\rho \\ \rho & -t \end{bmatrix} + \rho \det \begin{bmatrix} 1 & -t \\ \rho & -\rho \end{bmatrix} \\ &= -t(t^2 - \rho^2) - (-t + \rho^2) + \rho(-\rho + t\rho) = -t^3 + 3t + t - 3 - 3 + 3t \\ &= -t^3 + 7t - 6 = -(t-1)(t-2)(t+3). \end{aligned}$$

It follows that the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = -3$.

Example of orthogonal diagonalisation

$$\rho = \sqrt{3} \qquad A = \begin{bmatrix} 0 & 1 & \rho \\ 1 & 0 & -\rho \\ \rho & -\rho & 0 \end{bmatrix} \qquad \begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 2 \\ \lambda_3 &= -3. \end{aligned}$$

Eigenvectors can be found by row-reduction:

$$A - I = \begin{bmatrix} -1 & 1 & \rho \\ 1 & -1 & -\rho \\ \rho & -\rho & -1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & -\rho \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A - 2I = \begin{bmatrix} -2 & 1 & \rho \\ 1 & -2 & -\rho \\ \rho & -\rho & -2 \end{bmatrix} \to \begin{bmatrix} 1 & -2 & -\rho \\ 0 & -3 & -\rho \\ 0 & \rho & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -\rho/3 \\ 0 & 1 & \rho/3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A + 3I = \begin{bmatrix} 3 & 1 & \rho \\ 1 & 3 & -\rho \\ \rho & -\rho & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & -\rho \\ 0 & -8 & 4\rho \\ 0 & -4\rho & 6 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \rho/2 \\ 0 & 1 & -\rho/2 \\ 0 & 0 & 0 \end{bmatrix}$$

From this we can read off the following eigenvectors:

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} \rho/3 \\ -\rho/3 \\ 1 \end{bmatrix} \qquad u_3 = \begin{bmatrix} -\rho/2 \\ \rho/2 \\ 1 \end{bmatrix}.$$

Example of orthogonal diagonalisation

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 2 \\ \lambda_3 &= -3 \end{aligned} \qquad u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad u_2 = \begin{bmatrix} \rho/3 \\ -\rho/3 \\ 1 \end{bmatrix} \qquad u_3 = \begin{bmatrix} -\rho/2 \\ \rho/2 \\ 1 \end{bmatrix}$$

Because the matrix A is symmetric and the eigenvalues are distinct, it is automatic that the eigenvectors u_i are orthogonal to each other. However, they are not normalised: instead we have

$$\begin{aligned} u_1.u_1 &= 1^2 + 1^2 = 2 \\ u_2.u_2 &= (\rho/3)^2 + (-\rho/3)^2 + 1^2 = 1/3 + 1/3 + 1 = 5/3 \\ u_3.u_3 &= (-\rho/2)^2 + (\rho/2)^2 + 1^2 = 3/4 + 3/4 + 1 = 5/2. \end{aligned}$$

The vectors $v_i = u_i/\sqrt{u_i.u_i}$ form an orthonormal basis of eigenvectors. Explicitly, this works out as follows:

$$v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1/\sqrt{5} \\ -1/\sqrt{5} \\ \sqrt{3/5} \end{bmatrix} \qquad v_3 = \begin{bmatrix} -\sqrt{3/10} \\ \sqrt{3/10} \\ \sqrt{2/5} \end{bmatrix}.$$

Example of orthogonal diagonalisation

$$\begin{aligned} \lambda_1 &= 1 \\ \lambda_2 &= 2 \\ \lambda_3 &= -3 \end{aligned} \qquad v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1/\sqrt{5} \\ -1/\sqrt{5} \\ \sqrt{3/5} \end{bmatrix} \qquad v_3 = \begin{bmatrix} -\sqrt{3/10} \\ \sqrt{3/10} \\ \sqrt{2/5} \end{bmatrix}.$$

The eigenvectors v_i form orthonormal basis for $\mathbb{R}^3$.

It follows that if we put

$$U = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{5} & -\sqrt{3/10} \\ 1/\sqrt{2} & -1/\sqrt{5} & \sqrt{3/10} \\ 0 & \sqrt{3/5} & \sqrt{2/5} \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

then U is an orthogonal matrix and $A = UDU^T$.

Square roots of positive matrices

Corollary 23.18: Let A be an $n \times n$ real symmetric matrix, and suppose that all the eigenvalues of A are positive.
Then there is a real symmetric matrix B such that $A = B^2$.

Proof.
Choose an orthonormal basis of eigenvectors $u_1, \dots, u_n$, and let λ_i be the eigenvalue of u_i.
Put $U = [u_1|\cdots|u_n]$ and $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.
We saw in Corollary 23.15 that U is orthogonal (so $U^TU = I = UU^T$) and that $A = UDU^T$.

By assumption the eigenvectors λ_i are positive, so we have a real diagonal matrix $E = \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Put $B = UEU^T$. It is clear that $E^T = E$, and it follows that

$$B^T = (UEU^T)^T = U^{TT}E^TU^T = UEU^T = B.$$

We also have

$$B^2 = UEU^TUEU^T = UEEU^T = UDU^T = A.$$

□