Eigenvalues and Eigenvectors

Uday N. Gaitonde

IIT Bombay
IIT Dharwad

September 10, 2018

Study the spreadsheet.

Successive Multiplication

Our transition matrix was

$$\mathbf{A} = \begin{bmatrix} 0.40 & 0.60 \\ 0.80 & 0 \end{bmatrix}.$$

Repeated pre-multiplication led to vectors going as $\mathbf{x} \leftarrow 0.92\mathbf{x}$, where $\mathbf{x}$ is proportional to $\begin{bmatrix} 1.00 \\ 0.87 \end{bmatrix}$.

Formally, 0.92 is an **eigenvalue** of $\mathbf{A}$ and $\begin{bmatrix} 1.00 \\ 0.87 \end{bmatrix}$ is the corresponding **eigenvector**.

Formal Definitions

The formal defining relation is
$\mathbf{Ax} = \lambda\mathbf{x}$
where $\mathbf{A}$ is a square matrix $(n \times n)$
and $\mathbf{x}$ is a column vector of size n.

We can rewrite the equation as
$\mathbf{Ax} = \lambda\mathbf{Ix}$,
$\therefore \mathbf{Ax} - \lambda\mathbf{Ix} = \mathbf{0}$,
$\therefore (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$.

For a non-trivial solution to exist, we need
$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.
This is called the **characteristic equation** of $\mathbf{A}$.

Definitions (cont)

$\det(\mathbf{A} - \lambda\mathbf{I})$ is called the **characteristic determinant** of $\mathbf{A}$.

$(\mathbf{A} - \lambda\mathbf{I})$ is called the **characteristic matrix** of $\mathbf{A}$.

The characteristic equation of $\mathbf{A}$ can be expanded as an n-th degree polynomial.
This polynomial is called the **characteristic polynomial** of $\mathbf{A}$.

Example

In our case, $\mathbf{A} = \begin{bmatrix} 0.40 & 0.60 \\ 0.80 & 0 \end{bmatrix}$.

The characteristic matrix is $\begin{bmatrix} 0.40-\lambda & 0.60 \\ 0.80 & 0-\lambda \end{bmatrix}$.

The characteristic determinant is $\begin{vmatrix} 0.40-\lambda & 0.60 \\ 0.80 & 0-\lambda \end{vmatrix}$.

The characteristic equation is
$(0.40-\lambda)(0-\lambda) - (0.80)(0.60) = 0$,
i.e. $\lambda^2 - 0.4\lambda - 0.48 = 0$.

The LHS of this equation is the characteristic polynomial.

Example (cont)

The solution of the characteristic equation is:

$\lambda_1 = 0.9211102\ldots, \quad \lambda_2 = -0.5211102\ldots.$

These are the **eigenvalues** of the matrix $\mathbf{A}$.

Question: Why did not notice the presence of λ_2 in the evolution?

Example (cont)

To determine eigenvectors we need to solve:

$$\begin{array}{rcrcl} (0.40-\lambda)x_1 & + & (0.60)x_2 & = & 0 \\ (0.80)x_1 & + & (-\lambda)x_2 & = & 0 \end{array}$$

once with $\lambda = \lambda_1$ and then with $\lambda = \lambda_2$.
Since the equations will be homogeneous, the eigenvectors will have one component arbitrary.

Example (cont)

Let us take the second component as 1 (arbitrarily), for either eigenvector.

Then we have, the eigenvector for λ_1: $\mathbf{x}_{(1)} = \begin{bmatrix} 1.15139 \\ 1 \end{bmatrix}$

and the eigenvector for λ_2: $\mathbf{x}_{(2)} = \begin{bmatrix} -0.65139 \\ 1 \end{bmatrix}$

Note: Any eigenvector can be scaled, and will still remain an eigenvector.

Example (cont)

We may **normalise** an eigenvector, by scaling it such its length, or Euclidean norm, is 1.
In our case, the normalised eigenvectors will be:

$$\mathbf{x}_{(1)} = \begin{bmatrix} 0.75500 \\ 0.65573 \end{bmatrix}$$

$$\mathbf{x}_{(2)} = \begin{bmatrix} -0.65573 \\ 0.75500 \end{bmatrix}$$

Eigenvalues

We expect, generally, the matrix $\mathbf{A}$ to be non-zero.

Hence, the characteristic polynomial will be non-trivial.

Thus, for an $n \times n$ matrix, we will have at least one eigenvalue, and at most n numerically different eigenvalues.

Eigenvalues and Eigenvectors

Generally, for a given matrix, eigenvalues are determined first, and then
eigenvectors are computed for each eigenvalue.
For a given eigenvalue λ, the eigenvector $\mathbf{x}$ is obtained by solving $(\mathbf{A} - \mathbf{I}\lambda)\mathbf{x} = \mathbf{0}$.

Eigenvalues and Eigenvectors (cont)

Since the determinant

$$\begin{vmatrix} (a_{11}-\lambda) & a_{12} & \cdots & a_{1n} \\ a_{21} & (a_{22}-\lambda) & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & (a_{nn}-\lambda) \end{vmatrix} = 0,$$

there may not be a unique solution.

Eigenvalues and Eigenvectors (cont)

Recommended procedure:

Reduce the matrix to a row-echelon form.

At least one row at the bottom will be all zeros.

That makes x_n arbitrary.

Select any nonzero value for x_n, and
solve for $x_{n-1}, x_{n-2}, \ldots, x_1$.

This gives us an eigenvector corresponding to the λ chosen.

Eigenvalues and Eigenvectors (cont)

The n eigenvalues of an $n \times n$ matrix may not all be distinct.

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

We have: $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 1$.

Two eigenvalues are equal.

The number of equal eigenvalues is called the **algebraic multiplicity** of that eigenvalue, as is denoted by M_λ.

In this case, $M_{\lambda_1} = 1, M_{\lambda_2} = M_{\lambda_3} = 2$.

Eigenvalues and Eigenvectors (cont)

Determination of eigenvectors. For $\lambda_1 = 2$:

$$\begin{array}{rcrcrcll} (2-2)x_1 & + & 0x_2 & + & 0x_3 & = & 0 & (*) \\ 0x_1 & + & (1-2)x_2 & + & 0x_3 & = & 0 & \\ 0x_1 & + & 0x_2 & + & (1-2)x_3 & = & 0 & \end{array}$$

$(*) \implies x_1$ may have any value (say 1).
$x_2 = 0,\ x_3 = 0$.

Thus, $\mathbf{x}_{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

This is normalised.

Eigenvalues and Eigenvectors (cont)

For $\lambda_2 = \lambda_3 = 1$:

$$\begin{array}{rcrcrcll} (2-1)x_1 & + & 0x_2 & + & 0x_3 & = & 0 & \\ 0x_1 & + & (1-1)x_2 & + & 0x_3 & = & 0 & (*) \\ 0x_1 & + & 0x_2 & + & (1-1)x_3 & = & 0 & (*) \end{array}$$

Thus, $x_1 = 0$.
$(*) \implies x_2$ and x_3 may have any values.
$x_2 = 0, \; x_3 = 0$.

Selecting $(x_2, x_3) = (1, 0)$ and then $(0, 1)$, we get

$$\mathbf{x}_{(2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{x}_{(3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

These are linearly independent (and also normalised).

Eigenvalues and Eigenvectors (cont)

$\mathbf{x}_{(2)}$ and $\mathbf{x}_{(3)}$ are **two linearly independent** eigenvectors of $\lambda_2 = \lambda_3 = 1$.

The number of linearly independent eigenvectors of a multiple eigenvalue is called the **geometric multiplicity** of that eigenvalue, and is denoted by m_λ.

In this case, $m_{\lambda_1} = 1, m_{\lambda_2} = m_{\lambda_3} = 2$.

Eigenvalues and Eigenvectors (cont)

Some relations:

$$\sum_{\text{distinct } \lambda\text{s}} M_\lambda = n$$

Also

$$M_\lambda \geq m_\lambda$$

$\Delta_\lambda = (M_\lambda - m_\lambda)$ is known as the **defect** of that λ.

The defect is always zero or positive.[1]

[1] Study the example of a *defect*ive eigenvalue from Kreyszig.

A Vector Space

Let $\mathbf{x}$ and $\mathbf{y}$ be linearly independent eigenvectors of the same eigenvalue λ of a matrix $\mathbf{A}$.
Then, given $\alpha, \beta \in \mathbb{R}, \alpha \neq 0, \beta \neq 0$, the following will also be eigenvectors of that λ:

$$\begin{gathered} \alpha\mathbf{x} \\ \beta\mathbf{y} \\ \alpha\mathbf{x} + \beta\mathbf{y} \end{gathered}$$

Thus $\mathbf{x}$ and $\mathbf{y}$ together form the basis of a vector space, in which each member (vector) is an eigenvector of $\mathbf{A}$ for the given λ.

For a Real Matrix

Let $\mathbf{A}$ be a real matrix (all elements real).
$\implies$ All coefficients the characteristic polynomial are real.
$\implies$ All eigenvalues (roots of the CP) are either real or complex (in conjugate pairs).
If any λ is complex, then its eigenvector will also be complex.

Any odd-sized matrix will have at least one real eigenvalue.

A Theorem

A matrix $\mathbf{A}$ and its transpose $\mathbf{A}^T$ have the same set of eigenvalues.

Reason: transposing a determinant does not change its value.

Orthogonal Matrices

A matrix $\mathbf{A}$ is symmetric iff $\mathbf{A}^T = \mathbf{A}$.

A matrix $\mathbf{A}$ is skew-symmetric iff $\mathbf{A}^T = -\mathbf{A}$.

A matrix $\mathbf{A}$ is **orthogonal** iff $\mathbf{A}^T = \mathbf{A}^{-1}$.

Examples:

$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ is symmetric.

$\mathbf{B} = \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix}$ is skew-symmetric.

$\mathbf{C} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

Check Orthogonality

$$
\begin{aligned}
\mathbf{C}\mathbf{C}^T &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \mathbf{I}
\end{aligned}
$$

$\implies$ $\mathbf{C}$ is orthogonal.

Eigenvalues of an Orthogonal Matrix

Theorem
The eigenvalues of any symmetric matrix are real.

Theorem
The eigenvalues of any skew-symmetric matrix are either purely imaginary, or zero.

Exercises:
Demonstrate these theorems by working out with 2×2 matrices.
Attempt to do the same with 3×3 matrices.

Orthogonal Transformations

Let $\mathbf{A}$ be an orthogonal matrix ($n \times n$).
Let $\mathbf{x}$ be a vector (column, n).

Then $\mathbf{y} = \mathbf{A}\mathbf{x}$ is called an **orthogonal transformation** of $\mathbf{x}$ into $\mathbf{y}$.

In 2D and 3D space, an orthogonal transformation represents a rotation and/or a reflection. (Also in n-dimensional space $\mathbb{R}^n$).

Orthogonal Transformations (cont)

Let $\mathbf{a}$ and $\mathbf{b}$ be two n-vectors (column).
Then $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b}$ is called the **inner product** of $\mathbf{a}$ and $\mathbf{b}$.
This is similar to the dot product of the two vectors.

Let $\mathbf{A}$ be any $n \times n$ **orthogonal** matrix.

Theorem
If $\mathbf{u} = \mathbf{A}\mathbf{a}$ and $\mathbf{v} = \mathbf{A}\mathbf{b}$, then $\mathbf{u} \cdot \mathbf{v} = \mathbf{a} \cdot \mathbf{b}$.

Thus, an orthogonal transformation preserves inner product.

The proof is simple.

Orthogonal Transformations (cont)

Proof

$$\begin{aligned}
\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} &= (\mathbf{Aa})^T(\mathbf{Ab}) && \text{(transformation)}\\
&= (\mathbf{a}^T\mathbf{A}^T)(\mathbf{Ab}) && \text{(expansion)}\\
&= \mathbf{a}^T(\mathbf{A}^T\mathbf{A})\mathbf{b} && \text{(association)}\\
&= \mathbf{a}^T(\mathbf{A}^{-1}\mathbf{A})\mathbf{b} && \text{(orthogonality)}\\
&= \mathbf{a}^T(\mathbf{I})\mathbf{b} && \text{(inverse)}\\
&= \mathbf{a}^T\mathbf{b} && \text{(identity)}\\
&= \mathbf{a}\cdot\mathbf{b}
\end{aligned}$$

Orthogonal Transformations (cont)

Corollary
Let $\mathbf{u} = \mathbf{A}\mathbf{a}$, where $\mathbf{A}$ is orthogonal.
Then $\mathbf{u} \cdot \mathbf{u} = \mathbf{a} \cdot \mathbf{a}$.
$\therefore |\mathbf{u}|^2 = |\mathbf{a}|^2 \implies |\mathbf{u}| = |\mathbf{a}|$.
$|\mathbf{u}|$ is the magnitude of the vector $\mathbf{u}$, often called the Euclidean norm, and denoted by $\|\mathbf{u}\|$.

Thus, the length/magnitude of a vector is preserved after an orthogonal transformation.

Orthogonal Matrices

Theorem
A real square matrix $\mathbf{A}, (n \times n)$ is orthogonal iff its column vectors $a_1, a_2, \ldots, a_n$ (and also its row vectors) form an **orthonormal** system, *i.e.*

$$\begin{aligned} \mathbf{a}_i \cdot \mathbf{a}_j &= 0 \quad \text{if } i \neq j, \\ &= 1 \quad \text{if } i = j. \end{aligned}$$

Proof
The proof requires a study of the matrix $\mathbf{Z} \equiv \mathbf{A}^T\mathbf{A}$ and the realisation that an element of $\mathbf{Z}$, z_{ij}, is the inner product of the vectors $\mathbf{a}_i$ and $\mathbf{a}_j$, either of which is a column vector of $\mathbf{A}$.

Orthogonal Matrices (cont)

Proof (cont)

If $\mathbf{A}$ is orthogonal, then $\mathbf{Z}$, by definition is a unit matrix, and this leads to the conclusion that the set of vectors $\mathbf{a}_i$, which represent the columns of $\mathbf{A}$, is an orthonormal set of vectors.

On the other hand, if the set of vectors $\mathbf{a}_i$, which represent the columns of $\mathbf{A}$, is an orthonormal set of vectors, then $\mathbf{Z}$ turns out to be a unit matrix, which leads to the conclusion that $\mathbf{A}$ is an orthogonal matrix.

Orthogonal Matrices (cont)

Theorem

The determinant of an orthogonal matrix is either $+1$ or -1.

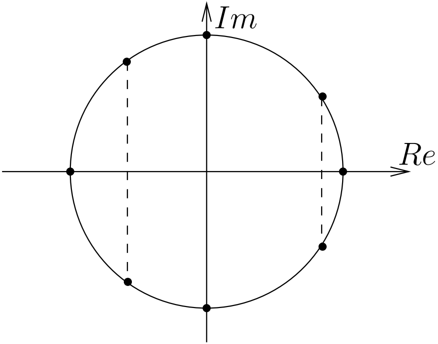
Proof

$\because \det \mathbf{A} = \det \mathbf{A}^T$, we have

$$\begin{array}{rcl} (\det \mathbf{A})^2 & = & (\det \mathbf{A})(\det \mathbf{A}) \\ & = & (\det \mathbf{A})(\det \mathbf{A}^T) \\ & = & \det(\mathbf{A}\mathbf{A}^T) \\ & = & \det(\mathbf{A}\mathbf{A}^{-1}) \\ & = & \det(\mathbf{I}) \\ & = & 1 \\ \therefore \det \mathbf{A} & = & +1 \text{ or } -1 \end{array}$$

Orthogonal Matrices (cont)

Theorem
Any eigenvalue of an orthogonal matrix has an absolute value of 1. That is the eigenvalues are $+1$ or -1 or occur in complex conjugate pairs.



Properties of Eigenvectors — Eigenbases

Eigenvectors of an $n \times n$ matrix $\mathbf{A}$ may or may not form a basis for n-dimensional real space, $\mathbb{R}^n$.

Suppose such a basis exists (*i.e.* when all eigenvectors are linearly independent), then any n-vector $\mathbf{x}$ can be represented as a linear combination of the eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.
$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots c_n\mathbf{x}_n$

We can then replace the multiplication of matrix $\mathbf{A}$ and any vector $\mathbf{x}$ by a linear combination of vectors.

Eigenbases (cont)

To demonstrate this, let the corresponding eigenvalues be $\lambda_1, \lambda_2, \ldots, \lambda_n$. (These need not be all distinct.)
Then we have $\mathbf{A}\mathbf{x}_j = \lambda_j \mathbf{x}_j$, and hence

$$\begin{aligned}
\mathbf{A}\mathbf{x} &= \mathbf{A}(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_n\mathbf{x}_n) \\
&= c_1(\mathbf{A}\mathbf{x}_1) + c_2(\mathbf{A}\mathbf{x}_2) + \ldots + c_n(\mathbf{A}\mathbf{x}_n) \\
&= c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \ldots + c_n\lambda_n\mathbf{x}_n
\end{aligned}$$

Thus, the multiplication-of-matrix-and-vector is replaced by scaled-sum-of-vectors.

Eigenbases (cont)

Theorem
If $\mathbf{A}$ $(n \times n)$ has n distinct eigenvalues, then its eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ form a basis for $\mathbb{R}^n$. (That is, the eigenvectors are all linearly independent.)

The proof is by *reductio-ad-absurdum*.[2]

The converse of this theorem is not true.

[2] Study the proof in Kreyszig.

Eigenbases (cont)

If the n eigenvalues are not distinct, then we need to look at the defects, if any.

If there is no defect, then we still have n distinct eigenvalues, and a basis is available.

But if a defect exists, then a basis may not exist.

Theorem
A symmetric matrix has an orthonormal basis for $\mathbb{R}^n$.

Similarity of Matrices

Let $\mathbf{A}$ be an $n \times n$ matrix.

Let $\widehat{\mathbf{A}}$ be another $n \times n$ matrix.

If an invertible matrix $\mathbf{P}$ exists, such that

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

then $\widehat{\mathbf{A}}$ is defined as similar to $\mathbf{A}$,
and the transformation $\widehat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is called a similarity transformation.

Similarity of Matrices (cont)

Theorem
If $\widehat{\mathbf{A}}$ is similar to $\mathbf{A}$, then
(eigenvalues of $\widehat{\mathbf{A}}$) $==$(eigenvalues of $\mathbf{A}$), and
if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ corresponding to some eigenvalue λ, then $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ is an eigenvector of $\widehat{\mathbf{A}}$ corresponding to the same eigenvalue λ.

Similarity of Matrices (cont)

Proof

$$
\begin{array}{rcll}
\mathbf{Ax} & = & \lambda\mathbf{x} & (\mathbf{x} \neq \mathbf{0}) \\
\therefore \mathbf{P}^{-1}\mathbf{Ax} & = & \mathbf{P}^{-1}\lambda\mathbf{x} & = \lambda(\mathbf{P}^{-1}\mathbf{x}) \\
\text{But } \mathbf{P}^{-1}\mathbf{Ax} & = & \mathbf{P}^{-1}\mathbf{AIx} & = \mathbf{P}^{-1}\mathbf{A}(\mathbf{PP}^{-1}\mathbf{x}) \\
 & = & (\mathbf{P}^{-1}\mathbf{AP})\mathbf{P}^{-1}\mathbf{x} & = \widehat{\mathbf{A}}(\mathbf{P}^{-1}\mathbf{x}) \\
\therefore \widehat{\mathbf{A}}(\mathbf{P}^{-1}\mathbf{x}) & = & \lambda(\mathbf{P}^{-1}\mathbf{x}) &
\end{array}
$$

$\implies \lambda$ is an eigenvalue of $\widehat{\mathbf{A}}$ and $\mathbf{P}^{-1}\mathbf{x}$ is the corresponding eigenvector.
(Note: $\mathbf{P}^{-1}\mathbf{x} \neq \mathbf{0}$. $\because$ if $\mathbf{P}^{-1}\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = \mathbf{Ix} = \mathbf{PP}^{-1}\mathbf{x} = \mathbf{P0} = \mathbf{0}$; but $\mathbf{x} \neq \mathbf{0}$.)

Diagonalisation of a Matrix

A diagonalised matrix of a given matrix $\mathbf{A}(n \times n)$ is a matrix similar to $\mathbf{A}$, which is diagonal, with each diagonal element an eigenvalue of $\mathbf{A}$.

Let $\mathbf{A}$ have a basis in its eigenvectors (*i.e.* its eigenvectors be all linearly independent). Let $\mathbf{X}$ be an $n \times n$ matrix with the eigenvectors of $\mathbf{A}$ as its column vectors. Then
$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$
is a diagonal matrix with its main diagonal populated by the eigenvalues of $\mathbf{A}$.
Also:
$\mathbf{D}^m = \mathbf{X}^{-1}\mathbf{A}^m\mathbf{X}$, where m is a positive integer.[3]

[3] Study and absorb the proof in Kreyszig.

Quadratic Form

Let $\mathbf{x}$ be an n-vector, and $\mathbf{A}$ an $n \times n$ matrix. Then

$$Q = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j x_k$$

is called a quadratic form in the components of $\mathbf{x}$ with $\mathbf{A}$ as its coefficient matrix.

Expanding the summations:

$$\begin{array}{ccccccccc} Q & = & a_{11}x_1^2 & + & a_{12}x_1x_2 & + & \cdots & + & a_{1n}x_1x_n \\ & + & a_{21}x_2x_1 & + & a_{22}x_2^2 & + & \cdots & + & a_{2n}x_2x_n \\ & & \vdots & & \vdots & & \ddots & & \vdots \\ & + & a_{n1}x_nx_1 & + & a_{n2}x_nx_2 & + & \cdots & + & a_{nn}x_n^2 \end{array}$$

Quadratic Form (cont)

One can always assume that $\mathbf{A}$ is symmetric.

Homework

Show that
$Q = \mathbf{x}^T\mathbf{A}\mathbf{x} = \mathbf{x}^T\mathbf{B}\mathbf{x}$
where $\mathbf{B}$ is the symmetric matrix $\mathbf{B} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T)$.
Also show that $(\mathbf{A} - \mathbf{A}^T)$ is skew-symmetric and does not contribute to Q.

Quadratic Form in Geometry

A common quadratic form in coordinate geometry pertains to the conic sections. The general form is:
$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$
$(A, \ldots, F$: real, not all zero.)
This can be written in a quadratic form as:

$$\underbrace{\begin{bmatrix} x & y & 1 \end{bmatrix}}_{\mathbf{x}^T} \underbrace{\begin{bmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{bmatrix}}_{\mathbf{A} \text{ (symmetric!)}} \underbrace{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}_{\mathbf{x}} = 0$$

Principal Axes in Quadratic Form

The coefficient matrix $\mathbf{A}$ in a quadratic form $Q = \mathbf{x}^T \mathbf{A} \mathbf{x}$ is symmetric.

Hence, $\mathbf{A}$ has an orthonormal basis of eigenvectors.

Hence, if we use these eigenvectors as column vectors and create a matrix $\mathbf{X}$, that matrix will be an orthogonal matrix.

Hence, $\mathbf{X}^T = \mathbf{X}^{-1}$.

Let $\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$ be the diagonalisation of $\mathbf{A}$.

$\therefore \mathbf{X}\mathbf{D}\mathbf{X}^{-1} = \mathbf{X}\mathbf{X}^{-1}\mathbf{A}\mathbf{X}\mathbf{X}^{-1} = \mathbf{A}$.

$\therefore \mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1} = \mathbf{X}\mathbf{D}\mathbf{X}^T$

Principal Axes in Quadratic Form (cont)

We have: $Q = \mathbf{x}^T\mathbf{A}\mathbf{x} = (\mathbf{x}^T\mathbf{X})\mathbf{D}(\mathbf{X}^T\mathbf{x})$.

Now, let $\mathbf{y} = \mathbf{X}^T\mathbf{x} = \mathbf{X}^{-1}\mathbf{x}$.
$\therefore \mathbf{X}\mathbf{y} = (\mathbf{X}\mathbf{X}^{-1})\mathbf{x} = \mathbf{x}$.

We also have: $\mathbf{x}^T\mathbf{X} = (\mathbf{X}^T\mathbf{x})^T = \mathbf{y}^T)$.

$$\begin{aligned} Q &= (\mathbf{x}^T\mathbf{X})\mathbf{D}(\mathbf{X}^T\mathbf{x}) \\ &= \mathbf{y}^T\mathbf{D}\mathbf{y}. \end{aligned}$$

This is an equivalent quadratic form with a diagonal coefficient matrix $\mathbf{D}$, wherein only the diagonal terms remain: $Q = \mathbf{y}^T\mathbf{D}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$.

This brings us to the ...

Principal Axes Theorem

Given $Q = \mathbf{x}^T\mathbf{A}\mathbf{x};\ \mathbf{A}$: symmetric, then

$\mathbf{x} = \mathbf{X}\mathbf{y}$ transforms it to the **canonical form**:

$Q = \mathbf{y}^T\mathbf{D}\mathbf{y}$,

where $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$,

where $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is the **spectrum** of $\mathbf{A}$, and

$\mathbf{X}$ is an orthogonal matrix with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ as column vectors.