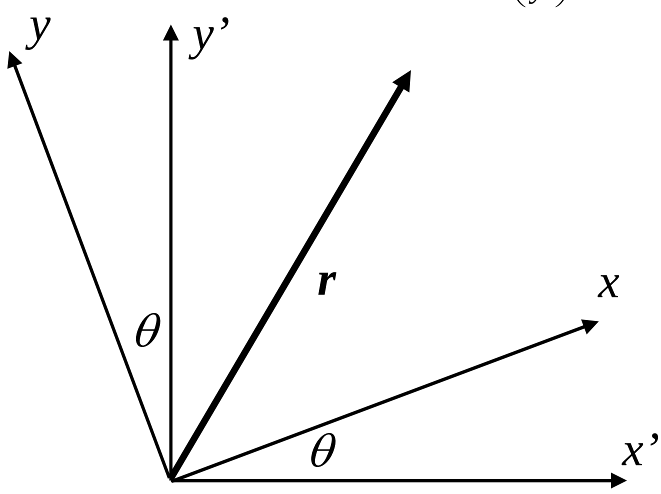
Fact Sheet 11 – Matrices II: Rotation

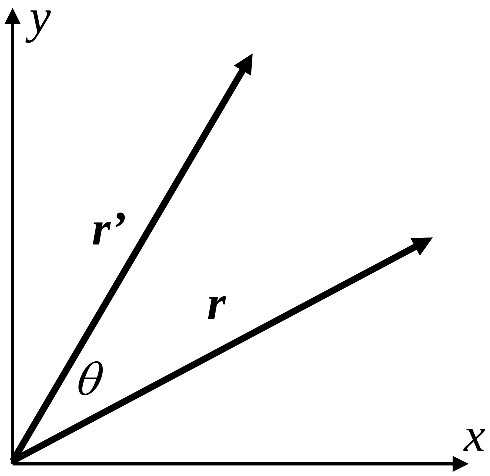
Consider a vector $\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ in the unprimed coordinate system, with coordinates $\mathbf{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ in the primed system, rotated clockwise by an angle θ with respect to the unprimed system, see diagram. It is easy to show that

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

which on matrix form reads $\mathbf{r}' = \mathbf{R}_\theta \mathbf{r}$.

The mathematics has been set up to transform the *coordinates* of a given vector from one set of axes *(x,y)* to another *(x',y')*.





However, an alternative interpretation is possible, namely that the rotation matrix is transforming one vector **r** into another **r'** against *fixed* axes. This second interpretation is shown in the diagram to the left. Note that according to this interpretation, the matrix $\mathbf{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ represents an <u>*anti-lockwise rotation of the vector*</u> **r** by an angle θ to create $\mathbf{r}'$ with fixed axis, whereas, in the first interpretation, the same matrix represents a <u>*clockwise rotation of the axes*</u> by an angle θ for a given vector.

In 3D, an *anti-clockwise rotation of the vectors* by an angle θ about the +*z*-axis is represented by a matrix of the form $\mathbf{R}_\theta^z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Remember that the +*z*-axis comes <u>*out of the paper in a right-handed system*</u> (see Fact Sheet 3), and that it is sensible to define the sense of rotation around the +z direction. Within this convention, the matrix given represents a *clockwise rotation of axes* for a given vector, or an anti-*clockwise rotation of a vector* with respect to fixed axes. By analogy, rotations about the *x*- and *y*-axes by an angle θ are of the form $\mathbf{R}_\theta^x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$ and $\mathbf{R}_\theta^y = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$, respectively.

Matrices II: The Eigenvalue problem

The *eigenvalue problem* is defined by the equation

$$\mathbf{Tv} = \lambda \mathbf{v}\,, \tag{1}$$

where **T** is a square matrix, and **v** is a column matrix. Hence, one is seeking vectors **v**, known as *eigenvectors,* that are unchanged in direction by **T**, but are simply scaled in magnitude by the factors λ, known as *eigenvalues*. Equation (1) can be written

$$\mathbf{Tv} = \lambda \mathbf{Iv} \tag{2}$$

or, by simple re-arrangement,

$$(\mathbf{T} - \lambda \mathbf{I})\mathbf{v} = \mathbf{0}\,, \tag{3}$$

where **I** is the identity matrix. Equation (3) is a homogeneous equation where the matrix of coefficients is $\mathbf{T} - \lambda \mathbf{I}$. The homogeneous equation has non-trivial solutions $\mathbf{v} \neq \mathbf{0}$ *if and only if* the determinant of the matrix of coefficients is zero, that is,

$$\det\left(\mathbf{T} - \lambda \mathbf{I}\right) = 0\,. \tag{4}$$

Equation (4) is the so-called *characteristic equation*. Note the Eqs (1)-(4) are valid in $\mathbb{R}^n$, $n \geq 2$.

In $\mathbb{R}^2$, that is, for a 2×2 system, Eq.(1) reads

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \tag{5}$$

and Eq.(3) reads

$$\begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0\,. \tag{6}$$

Equation (6) represents is a homogeneous linear equation (two equations with two unknowns):

$$\begin{aligned} (a_{11} - \lambda)x + a_{12}y &= 0 \\ a_{21}x + (a_{22} - \lambda)y &= 0. \end{aligned} \tag{7}$$

These equations have non-trivial solutions (in addition to the trivial solution $x = y = 0$) *if and only if* the determinant of the matrix of coefficients is zero, that is,

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) - a_{21}a_{12} = 0 \tag{8}$$

or in other words

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{21}a_{12}) = 0\,. \tag{9}$$

The two roots of this quadratic equation are

$$\lambda_{1,2} = \frac{(a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2 + 4a_{21}a_{12}}}{2} \tag{10}$$

and λ_1 and λ_2 are the two eigenvalues of **T**.

In the case treated in the Lecture, $\mathbf{T}=\begin{pmatrix}5 & -1\\ -1 & 5\end{pmatrix}$ and $\lambda_{1,2}=\frac{10\pm 2}{2}=\begin{cases}6\\4\end{cases}$. These eigenvalues must now be inserted in turn in Eq.(7) to find the associated eigenvectors. One obtains

$$\lambda_1=6; \qquad -x_1-y_1=0 \Leftrightarrow y_1=-x_1; \qquad \mathbf{v}_1=\begin{pmatrix}1\\ -1\end{pmatrix}; \qquad \hat{\mathbf{v}}_1=\begin{pmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{pmatrix}, \tag{11a}$$

$$\lambda_2=4; \qquad x_2-y_2=0 \Leftrightarrow y_2=x_2; \qquad \mathbf{v}_2=\begin{pmatrix}1\\ 1\end{pmatrix}; \qquad \hat{\mathbf{v}}_2=\begin{pmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{pmatrix}, \tag{11b}$$

where $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2$ gives the eigenvectors in normalised form. Note that Eq.(5) can be written

$$\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x_1\\ y_1\end{pmatrix}=\lambda_1\begin{pmatrix}x_1\\ y_1\end{pmatrix}, \tag{12a}$$

$$\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x_2\\ y_2\end{pmatrix}=\lambda_2\begin{pmatrix}x_2\\ y_2\end{pmatrix}. \tag{12b}$$

These last two equations can be combined to read

$$\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x_1 & x_2\\ y_1 & y_2\end{pmatrix}=\begin{pmatrix}\lambda_1 x_1 & \lambda_2 x_2\\ \lambda_1 y_1 & \lambda_2 y_2\end{pmatrix}=\begin{pmatrix}x_1 & x_2\\ y_1 & y_2\end{pmatrix}\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}. \tag{13}$$

Introducing the matrix of the eigenvectors $\mathbf{V}=\begin{pmatrix}x_1 & x_2\\ y_1 & y_2\end{pmatrix}$ and the <u>*diagonal*</u> matrix of the eigenvalues $\mathbf{S}=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$, Eq.(13) reads

$$\mathbf{TV} = \mathbf{VS}. \tag{14}$$

Multiplying both sides of Eq.(14) from the left by $\mathbf{V}^{-1}$, the inverse of **V,** yields the important result

$$\mathbf{V}^{-1}\mathbf{TV} = \mathbf{S}. \tag{15}$$

The matrices **T** and **S** are said to be "similar", and **V** is said to "diagonalise **T** in a *similarity transformation*". Note that Eq.(13) applies irrespective of whether the eigenvectors have been normalised before the formation of **V**. However, in the special case where **T** is <u>*symmetric*</u> (i.e., $a_{21}=a_{12}$ as in the numerical example given above), it can be shown that $\mathbf{v}_1$ and $\mathbf{v}_2$ are <u>*orthogonal*</u>. And if the vectors are also <u>*normalised,*</u> see Eq.(11), then **V** becomes an <u>*orthogonal matrix*</u>, in which case $\mathbf{V}^{-1}=\mathbf{V}^t$, the transpose of the matrix **V**. For instance, in the example above,

$\mathbf{V}=\begin{pmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$ and $\mathbf{V}^{-1}=\mathbf{V}^t=\begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$. In this case

$$\mathbf{V}^{-1}\mathbf{TV}=\begin{pmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}\begin{pmatrix}5 & -1\\ -1 & 5\end{pmatrix}\begin{pmatrix}1/\sqrt{2} & 1/\sqrt{2}\\ -1/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}=\begin{pmatrix}6 & 0\\ 0 & 4\end{pmatrix}=\mathbf{S}. \tag{16}$$

Note that for a 3×3 system, Eq.(9) becomes a cubic equation, for a 4×4 system, it's a quartic and so on. If you have to solve a cubic equation, you may be able to guess one of the roots, λ_1, then divide the cubic equation with the factor $(\lambda-\lambda_1)$, to obtain a quadratic equation and then solve that to identify the other two roots λ_2, λ_3.