SQEP Tutorial Session 6: T72S01

Relates to Knowledge & Skills: ALL, but especially 1.5, 1.9, 1.10

Maths: Tensors and index notation; Transformation of tensors to a different coordinate system (rotations); 2D rotated stresses, formulae for principal stresses and orientation; Eigenvalues and eigenvectors of real symmetric matrices: relevance to principal stresses and their orientation: the reason why principal axes exist.

Last Update: 3/1/13

Qu.: What is the definition of the derivative of a function?

$\frac{df}{dx}$ is the gradient of the graph of $f(x)$ plotted against x. Algebraically this is,

$$\frac{df}{dx} = LIM_{\delta x \to 0}\left[\frac{f(x+\delta x) - f(x)}{\delta x}\right] \qquad (1)$$

The partial derivative is just the same with the independent variables held fixed, e.g.,

$$\frac{\partial f(x,y,z)}{\partial x} = LIM_{\delta x \to 0}\left[\frac{f(x+\delta x, y, z) - f(x,y,z)}{\delta x}\right] \qquad (2)$$

Qu.: How is matrix multiplication done?

Matrix times 'vector': $(A)\bar{v}$ gives another vector, $\bar{u}$, whose components are given by,

$$u_i = \sum_j A_{ij} v_j \qquad (3)$$

Strictly, $\bar{u}$ and $\bar{v}$ are column matrices, but like everyone else I shall often loosely refer to column or row matrices as 'vectors' though this is not really correct (see below).

A matrix times a matrix gives another matrix, $(M) = (A)(B)$, whose components are given by,

$$M_{ij} = \sum_k A_{ik} B_{kj} \qquad (4)$$

Qu.: What is the summation convention?

The summation convention is that repeated indices are assumed to be summed over. This just means we can stop explicitly including all the $\sum$ signs. So, the above definitions of multiplication can be written,

$$u_i = A_{ij} v_j \text{ and } M_{ij} = A_{ik} B_{kj} \qquad (5)$$

Qu.: A tensor is just a matrix, right?

Wrong.

Qu.: But a vector is just a set of three numbers, right?

Wrong.

Qu.: OK, so what is a vector?

A vector is a mathematical object with a magnitude and an orientation. It can be visualised as an arrow. A vector field (e.g., the velocity of a moving fluid) consists of an arrow associated with every point in space.

Qu.: OK, but these arrows have components, which are just three numbers

True, but the components of a vector are specific to the coordinate system used. If we use a different coordinate system, the components of the vector change even though the vector itself is unchanged. So a vector does not equal three numbers, but it can be described by three numbers with respect to a specified coordinate system.

But the vector can also be described just as well by a different set of three numbers (components) with respect to a different coordinate system. So the vector clearly does not ***equal*** any set of three numbers.

Qu.: What is the defining feature of a vector?

The defining feature of a vector, expressed in terms of its components, is that the components transform between coordinate systems in a manner which is uniquely determined by the coordinate system. That is - all vectors transform in the same way between a given pair of coordinate systems.

Qu.: If we know the coordinates of a vector in one coordinate system, how do we find its coordinates with respect to another coordinate system?

We shall assume that all the coordinate systems are Cartesian coordinate systems with a common origin, so that they differ from each other only by a 3D rotation. Suppose the coordinates of a point change from (x,y,z) in the original coordinate system to (x',y',z') in the rotated coordinate system. The two sets of coordinates are related by a matrix Q which represents the rotation,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = (Q)\begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{6}$$

where matrix multiplication is understood. But a rotation leaves the distance of the point from the origin unchanged, and this is given by Pythagoras in both sets of coordinates. Hence,

$$x'^2 + y'^2 + z'^2 = \begin{pmatrix} x & y & z \end{pmatrix}(Q)^T(Q)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2 \tag{7}$$

where ^{T} represents the transpose of the matrix. But this must hold for an arbitrary choice of the point (x,y,z), so that $(Q)^T(Q)$ must be just the unit matrix, i.e.,

$$(Q)^T(Q) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv \mathrm{I} \tag{8}$$

But the inverse of a matrix is defined to be such that $(Q)^{-1}(Q) = \mathrm{I}$, so that rotations must be represented by matrices whose inverse equals their transpose, $(Q)^{-1} \equiv (Q)^T$. Such matrices are called *orthogonal*.

The Q-matrix is given explicitly in terms of the unit vectors of the rotated coordinate axes. Suppose that $\hat{x}'$ is the row-matrix ('vector') made up of the components with respect to the (x,y,z) system of the unit vector in the rotated x' direction, and similarly for $\hat{y}'$ and $\hat{z}'$. The rotation matrix is given explicitly by,

$$(Q) = \begin{pmatrix} \hat{x}' \\ \hat{y}' \\ \hat{z}' \end{pmatrix} = \begin{pmatrix} \cos\theta_{x'x} & \cos\theta_{x'y} & \cos\theta_{x'z} \\ \cos\theta_{y'x} & \cos\theta_{y'y} & \cos\theta_{y'z} \\ \cos\theta_{z'x} & \cos\theta_{z'y} & \cos\theta_{z'z} \end{pmatrix} \quad (9)$$

where, for example, $\theta_{x'y}$ is the angle between the rotated x axis and the original y axis, etc. Of course, only three of these nine angles are independent. Note that the orthogonality of (Q) follows from $\hat{x}'^T\hat{x}' = 1$, $\hat{x}'^T\hat{y}' = 0$, etc.

Qu.: What is the 3D rotation matrix in terms of spherical polar angles?

Defining spherical polar angles θ and ϕ in the conventional manner, we may define our rotated Cartesian coordinate system (x',y',z') so that these rotated axes align with the polar directions $\hat{\theta}, \hat{\phi}, \hat{r}$ respectively. [Note that this order of correspondence means that for $\theta = 0, \phi = 0$, i.e., when no rotation is carried out, (x',y',z') becomes (x,y,z)]. In this case we get, for any vector $\bar{v}$,

$$\begin{pmatrix} v'_x \\ v'_y \\ v'_z \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \quad (9a)$$

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} v'_x \\ v'_y \\ v'_z \end{pmatrix} \quad (9b)$$

As an exercise the reader may like to check that the above matrices are orthogonal.

Qu.: So what is the strict definition of a vector in terms of its components?

A vector can be defined via its components by requiring that the components transform under rotations in the manner given by (6) when (Q) is given by (9). Thus a vector is defined by the transformation law,

$$v'_i = Q_{ij} v_j \quad (10)$$

(Note the summation convention).

Hence, *a set of three numbers defined with respect to a coordinate system, and varying depending on the coordinate system, can be recognised as the components of a vector if and only if the components transform between coordinate systems in the manner indicated by (10).*

What this really means is that there is an underlying object (the "arrow") which is completely independent of the coordinate system.

Qu.: What is a non-Euclidean, or non-Cartesian, tensor?

We are confining attention to transformations, (Q), which are rotations between Cartesian coordinate systems in an underlying Euclidean space. So we are considering only Cartesian (or Euclidean) vectors and tensors. Transformations between curvilinear coordinate systems provide a generalisation and give rise to non-Cartesian vectors and tensors. If the underlying space is curved they will also be non-Euclidean. We'll not go there.

Qu.: So, what is a tensor?

A tensor of rank r is described by a set of components with r subscripts, each of which transforms under rotations like a vector. For example, a second rank tensor, T_{ij} , is defined to transform as,

$$T'_{ij} = Q_{ik} Q_{jn} T_{kn} \tag{11}$$

In matrix notation this can be written,

$$(T') = (Q)(T)(Q)^T = (Q)(T)(Q)^{-1} \tag{12}$$

which is sometimes called a *similarity transformation*. A fourth rank tensor obeys,

$$C'_{ijkl} = Q_{ip} Q_{jq} Q_{km} Q_{\ln} C_{pqmn} \tag{13}$$

Note that matrix notation is useless for rank 3 tensors and higher.

Stress and strain are examples of second rank tensors.

An example of a fourth rank tensor is the elastic modulus relation between stress and strain: $\varepsilon_{ij} = C_{ijkl}\sigma_{kl}$. For anisotropic materials, e.g., single crystals, C_{ijkl} is more complicated than for isotropic materials. In a coordinate system not aligned with the crystal axes, Equ.(13) shows how to find the elastic moduli in the rotated system.

A vector is just a rank 1 tensor.

Qu.: Isn't this all a bit difficult?

All this difficulty is caused by the arbitrariness of the coordinate system we use to define the components of the tensor. There is an underlying invariant object, the tensor itself, which does not depend upon a coordinate system at all. But to do calculations we need to describe the tensor in terms of numbers, i.e., coordinates with respect to a coordinate system. The price we pay for this is that the coordinates have a large degree of arbitrariness about them. Tensor algebra is really about how to distinguish these arbitrary, coordinate dependent, features from the underlying invariant object.

[Aside: This, incidentally, is why the Theory of Relativity is the most inappropriately named theory in physics. It uses tensors as a means of expressing the underlying objective reality of physical quantities, independent of the "coordinate system", which, in spacetime, includes the state of motion of the observer. The true maxim of relativity theory is that "everything is objective" not relative. It is only the components specified in arbitrary coordinate systems (observers) which are relative].

Qu.: How do stresses (or strains) transform under rotation?

This is just a particular example of (12), where (Q) is given by (9). It is a bit horrid algebraically in 3D. But in 2D the rotation matrix reduces to simply,

$$(Q) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{14}$$

From (12) we get the rotated stresses to be,

$$\begin{pmatrix} \sigma'_x & \tau' \\ \tau' & \sigma'_y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \sigma_x & \tau \\ \tau & \sigma_y \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{15}$$

Carrying out the matrix multiplication, this reduces to,

$$\sigma'_x = \sigma_x \cos^2\theta + \sigma_y \sin^2\theta + \tau \sin 2\theta \tag{16a}$$

$$\sigma'_y = \sigma_y \cos^2\theta + \sigma_x \sin^2\theta - \tau \sin 2\theta \tag{16b}$$

$$\tau' = \tau \cos 2\theta + \frac{1}{2}(\sigma_y - \sigma_x)\sin 2\theta \tag{16c}$$

where we have also used the trigonometric identities $\sin 2\theta = 2\sin\theta\cos\theta$, $\cos 2\theta = \cos^2\theta - \sin^2\theta$.

Qu.: What does (16) tell us about the principal stresses in 2D?

The principal axes are the orientation of the coordinate system which results in no shear stresses. Thus, for a 2D problem, (16c) tells us that the orientation of the principal axes must be at an angle, θ, such that,

$$\tan 2\theta = \frac{2\tau}{\sigma_x - \sigma_y} \tag{17}$$

Qu.: What are the eigenvalues of a matrix?

In general, when a matrix multiplies a 'vector' (by which I really mean a column matrix) you just get a completely different 'vector'. However, for certain special vectors, the result equals the original vector times some number, i.e.,

$$(M)\bar{v} = \lambda\bar{v} \tag{18}$$

When (18) holds, $\bar{v}$ is said to be an eigenvector of (M) and λ the corresponding eigenvalue. In other words, the matrix multiplies an eigenvector just as if it were the number λ.

Note that the eigenvectors are only determined up to a normalisation constant. That is, if $\bar{v}$ satisfies (18) then so does $\mu\bar{v}$ for any number μ. From hereon we shall assume normalised eigenvectors, such that $\bar{v}^T\bar{v} = 1$.

Qu.: What's special about the eigenvectors of real symmetric matrices?

The eigenvectors of any real symmetric matrix are mutually orthogonal. This means that the eigenvectors of any real symmetric matrix form a Cartesian coordinate system. The proof follows. It is just as easy to carry out the proof for the more general case of a complex Hermitian matrix. An Hermitian matrix obeys $H^+ = H$, where the symbol $^+$ denotes the Hermitian conjugate, which is the combination of the complex conjugate and the transpose: $H^+ \equiv H^{*T}$. Hence, a real Hermitian matrix is symmetric.

Qu. What's special about the eigenvalues and eigenvectors of Hermitian matrices?

The eigenvectors of an Hermitian matrix are mutually orthogonal and its eigenvalues are real.

Proof: Let the eigenvectors and eigenvalues be respectively $\bar{v}_i$ and λ_i, where the subscript $_i$ labels the different eigenvectors/values. Hence, $H\bar{v}_i = \lambda_i \bar{v}_i$. Multiplying from the left by the Hermitian conjugate of eigenvector $\bar{v}_j$ gives $\bar{v}_j^{\,+} H\bar{v}_i = \lambda_i \bar{v}_j^{\,+} \bar{v}_i$. But we also have $H\bar{v}_j = \lambda_j \bar{v}_j$, and taking the Hermitian conjugate of both sides gives $\bar{v}_j^{+} H^{+} = \lambda_j^{*} \bar{v}_j^{+}$. Multiplying by $\bar{v}_i$ on the right gives $\bar{v}_j^{+} H^{+} \bar{v}_i = \bar{v}_j^{+} H\bar{v}_i = \lambda_j^{*} \bar{v}_j^{+} \bar{v}_i$, where we have used $H^{+} = H$. So we have shown that $\bar{v}_j^{\,+} H\bar{v}_i = \lambda_i \bar{v}_j^{\,+} \bar{v}_i = \lambda_j^{*} \bar{v}_j^{+} \bar{v}_i$. Now if we take $i = j$ then $\bar{v}_j^{+} \bar{v}_i \to \bar{v}_i^{+} \bar{v}_i \neq 0$ because it is the sum of absolute squares. So the fact that $\lambda_i \bar{v}_i^{\,+} \bar{v}_i = \lambda_i^{*} \bar{v}_i^{\,+} \bar{v}_i$ allows us to conclude that $\lambda_i = \lambda_i^{*}$, i.e., the eigenvalues are all real. We thus have $\lambda_i \bar{v}_j^{\,+} \bar{v}_i = \lambda_j \bar{v}_j^{+} \bar{v}_i$. In the case that the eigenvalues are different, $\lambda_i \neq \lambda_j$, this can only be true if $\bar{v}_j^{\,+} \bar{v}_i = 0$, i.e., the eigenvectors are orthogonal. QED.[1]

Qu.: What is meant by "putting a matrix in diagonal form"?

"Diagonal form" just means that the only non-zero elements of a matrix are on the principal diagonal. Hermitian matrices can be put in diagonal form by an orthogonal similarity transform, i.e., by a rotation. The rotation matrix in question is formed from the eigenvectors, i.e., in an arbitrary number of dimensions...

$$(Q)^{+} = \begin{pmatrix} \bar{v}_1 & \bar{v}_2 & \bar{v}_3 & etc. \end{pmatrix} \tag{19}$$

So that,

$$\begin{aligned}(Q)H(Q)^{+} &= \begin{pmatrix} \bar{v}_1^{+} \\ \bar{v}_2^{+} \\ \bar{v}_3^{+} \\ etc. \end{pmatrix} H \begin{pmatrix} \bar{v}_1 & \bar{v}_2 & \bar{v}_3 & etc. \end{pmatrix} = \begin{pmatrix} \bar{v}_1^{+} \\ \bar{v}_2^{+} \\ \bar{v}_3^{+} \\ etc. \end{pmatrix} \begin{pmatrix} \lambda_1 \bar{v}_1 & \lambda_2 \bar{v}_2 & \lambda_3 \bar{v}_3 & etc. \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 & 0 & \\ 0 & \lambda_2 & 0 & \\ 0 & 0 & \lambda_3 & \\ & & & etc. \end{pmatrix}\end{aligned} \tag{20}$$

The latter follows from orthonormality, $\bar{v}_j^{\,+} \bar{v}_i = \delta_{ij}$.

Qu.: What's the relevance of this for stress and strain?

Stress is a real symmetric matrix, and hence a special case of an Hermitian matrix. If a stress matrix is in diagonal form this just means that there are no shear terms. In other words, the diagonal form is obtained for the principal axes, and the diagonal terms are

[1] The alert will spot a loophole. It can happen that two or more distinct eigenvectors share the same eigenvalue. Such an eigenvalue is called *degenerate*. However, this does not really compromise the above picture. A set of mutually orthogonal eigenvectors always exists, but is no longer unique. Eigenvectors which belong to the same degenerate eigenvalue span a subspace within which any vector is an eigenvector.

the principal stresses. The above mathematics proves that there exists a rotation matrix, given in terms of the eigenvectors of the stress matrix by (19), which diagonalises the stress matrix. In other words a principal coordinate system always exists and the principal axes are just the eigenvectors of the stress matrix, and the principal stresses are its eigenvalues.

Note that the existence of a principal coordinate system follows from the symmetry of the stress matrix, and this in turn follows from moment equilibrium. Thus, the requirement for zero total moment requires the existence of a principal coordinate system. Strange but true.

The strain matrix is also symmetric and hence also must have a principal coordinate system in which the shear strains are zero.

Qu.: Are the principal coordinate systems for stress and strain the same?

For an isotropic elastic medium, yes. This follows simply from the fact that the shear stresses and shear strains are proportional, $G\gamma_{xy} = \sigma_{xy}$, etc. So if the shear stresses vanish in a particular coordinate system, so must the shear strains. However, for a general anisotropic medium, whilst principal stress and principal strain axes will always exist (due to the symmetry of the tensors) that will not necessarily be the same axes. The same follows in plasticity, since stress and strain are no longer proportional.

Qu.: How do you actually calculate the eigenvalues of a matrix?

Re-arranging $(M)\bar{v} = \lambda \bar{v}$ gives $(M - \lambda \mathrm{I})\bar{v} = \bar{0}$. But a set of equations $(A)\bar{v} = \bar{0}$ has a non-trivial solution only if $\|A\| = 0$, i.e., $\|M - \lambda \mathrm{I}\| = 0$. This is called the *secular equation*. In the general, 3D, case it means that the principal stresses are the three solutions to the cubic equation,

$$\left\| \begin{pmatrix} \sigma_x - \lambda & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y - \lambda & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z - \lambda \end{pmatrix} \right\| = 0 \qquad (21)$$

In practice, a general cubic is a bit of a pain (though the solution is given in http://rickbradford.co.uk/SolutionofCubic.pdf). The use of software which can extract eigenvalues automatically is advisable, e.g., Matlab. However, in 2D plane stress or plane strain it is easy. The secular equation reduces to,

$$\left\| \begin{pmatrix} \sigma_x - \lambda & \sigma_{xy} & 0 \\ \sigma_{xy} & \sigma_y - \lambda & 0 \\ 0 & 0 & \sigma_z - \lambda \end{pmatrix} \right\| = 0 \qquad (22)$$

So one principal stress is the out-of-plane stress, σ_z. The other two are found from $(\sigma_x - \lambda)(\sigma_y - \lambda) - \sigma_{xy}^2 = 0$ which gives the two remaining principal stresses to be,

$$\lambda_{\pm} = \frac{1}{2}\left\{ \sigma_x + \sigma_y \pm \sqrt{(\sigma_x - \sigma_y)^2 + 4\sigma_{xy}^2} \right\} \qquad (23)$$

Qu.: What is the comma notation?

A compact way of writing derivatives with respect to a coordinate is to define, say,

$$f_{,i} \equiv \frac{\partial f}{\partial x_i} \tag{24}$$

The nine possible spatial derivatives of the components v_i of a vector can therefore be written,

$$v_{i,j} \equiv \frac{\partial v_i}{\partial x_j} \tag{25}$$

For Cartesian tensors, the divergence of a vector can therefore be written,

$$\overline{\nabla} \cdot \overline{v} = \sum_{i=1}^{3} \frac{\partial v_i}{\partial x_i} \equiv v_{i,i} \tag{26}$$

where the last expression follows because we are also employing the summation convention over repeated indices.