DIAGONALIZATION OF THE STRESS TENSOR

INTRODUCTION

By the use of Cauchy's theorem we are able to reduce the number of stress components in the stress tensor to only nine values. An additional simplification of the stress can be achieved through diagonalization of the stress tensor. Diagonalization of the stress tensor reduces the number of components to only three. Many square matrices can be diagonalized.

In this simplified (diagonalized) version of the stress tensor, the principal planes have no stress along them and the principal axes are the only directions along which we have any stress. The principal axes directions are orthogonal to the principal planes.

That is, we start with a general matrix:

$$\sigma_{ij} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$$

And end with a simpler matrix:

$$\boldsymbol{\sigma}_{ij} = \begin{pmatrix} \boldsymbol{\sigma}_{11} & 0 & 0 \\ 0 & \boldsymbol{\sigma}_{22} & 0 \\ 0 & 0 & \boldsymbol{\sigma}_{33} \end{pmatrix}$$

The last matrix has been diagonalized. The resultant matrix is easier to handle. For example, the square of the diagonalized matrix is simply the square of each of the components.

> Recall Cauchy's Theorem that states:
>
> $$\boldsymbol{T}_i = \sigma_{ji}\boldsymbol{n}_j$$
>
> Where $\boldsymbol{T}$ is the traction/stress vector at a point on a plane with normal vector $\boldsymbol{n}$
>
> $\boldsymbol{T}$ the stress tensor is symmetric.

Definition

A square matrix , $\boldsymbol{M}$, can become diagonalized into another matrix, $\boldsymbol{D}$ by deriving $\boldsymbol{C}$, so that

$$\boldsymbol{D} = \boldsymbol{C}^{-1}\boldsymbol{M}\boldsymbol{C}$$

In the case that $\boldsymbol{M}$ is the stress tensor, $\boldsymbol{D}$ becomes a description of the same stress field from the perspective of a new, rotated co-ordinate system. From the point of view of this new stress matrix $\boldsymbol{M}$ is the stress described in the old co-ordinate system. So, in different words diagonalization gives the components of stress in a new, rotated coordinate system.

That is, we start with a general matrix:

$$\boldsymbol{\sigma}_{ij} = \begin{pmatrix} \boldsymbol{\sigma}_{11} & \boldsymbol{\sigma}_{12} & \boldsymbol{\sigma}_{13} \\ \boldsymbol{\sigma}_{21} & \boldsymbol{\sigma}_{22} & \boldsymbol{\sigma}_{23} \\ \boldsymbol{\sigma}_{31} & \boldsymbol{\sigma}_{32} & \boldsymbol{\sigma}_{33} \end{pmatrix}$$

and end with a simpler matrix:

$$\boldsymbol{\sigma}_{ij} = \begin{pmatrix} \boldsymbol{\sigma}_{11} & 0 & 0 \\ 0 & \boldsymbol{\sigma}_{22} & 0 \\ 0 & 0 & \boldsymbol{\sigma}_{33} \end{pmatrix}$$

Components are not the same before and after the diagonalization. The variable name is the same but the values are different.

When a matrix is diagonalizable, it means that the three basis vectors in the new Cartesian coordinate system are parallel (in the general case) to three non-basis traction vectors in the old coordinate system. These special three vectors formerly in the old coordinate system have new components in the new coordinate system. These new components are the rows of the new (diagonalized) stress tensor. Diagonalization requires us to find these vectors.

Eigenvalues and Eigenvectors

In other words, this means that:

$$\hat{x}_1' \parallel \boldsymbol{U} \text{ and } \hat{x}_2' \parallel \boldsymbol{V} \text{ and } \hat{x}_3' \parallel \boldsymbol{W}$$

where the basis vectors are in the new coordinate system (primes) and are parallel to vectors that were formerly in the old coordinate system. The new basis vectors and their corresponding former selves are called eigenvectors. We can express this another way:

$$\boldsymbol{U}' = \mu_1 \boldsymbol{U} \,,$$

$$\boldsymbol{V}' = \mu_2 \boldsymbol{V},$$

$\boldsymbol{W}' = \mu_3 \boldsymbol{W}$ where μ_1, μ_1, and μ_3, are constants, known as eigenvalues.

Let's take a simple example and consider only vectors in a 2-coordinate system that are acted upon by some matrix that we wish to diagonalize (our target).

$$\begin{pmatrix} \boldsymbol{U}_1' \\ \boldsymbol{U}_2' \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_2 \end{pmatrix}$$

Also,

$$\begin{pmatrix} \boldsymbol{U}_1' \\ \boldsymbol{U}_2' \end{pmatrix} = \mu_1 \begin{pmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_2 \end{pmatrix}$$

The simultaneous solution to this problem is given by Cramer's Rule:

$$\begin{aligned} 5U_1 - 2U_2 &= \mu_1 U_1 \\ -2U_1 + 2U_2 &= \mu_1 U_1 \end{aligned}$$

Now,

$$\begin{aligned} (5-\mu_1)U_1 - 2U_2 &= 0 \\ -2U_1 + (2-\mu_1)U_2 &= 0 \end{aligned}$$

So,

$\begin{vmatrix} 5-\mu_1 & -2 \\ -2 & 2-\mu_1 \end{vmatrix} = 0$ to obtain solutions to $\boldsymbol{U}_1$

This is also known as the <u>characteristic equation</u> of matrix $\boldsymbol{M}$

You should work out that μ_1 can have two values: 6 or 1

When μ_1=1 then we have

$$\begin{vmatrix} 5-1 & -2 \\ -2 & 2-1 \end{vmatrix} = 0$$

These are

DIAGONALIZATION OF THE STRAIN TENSOR, AN EXAMPLE

We will look at the diagonalization of strain instead of the case of stress as this will lead us in to the next section.

How to find the eigenvalues and eigenvectors of a symmetric 2x2 matrix

Introduction

We will leave the theoretical development of eigensystems for you to read in textbooks on linear algebra or tensor mathematics, or from reliable sources on the web such as those listed in the references section at the end of this document. Here, we accept that for any given stress or strain tensor, a coordinate system can be identified in which all of the shear stresses or shear strains have zero value, and the only non-zero values are along the diagonal of the tensor matrix. The eigenvectors are parallel to those special coordinate axes, and the eigenvalues are the values along the diagonal. Another way of characterizing them is that the eigenvectors are along the principal directions of the stress or strain ellipsoids, and the eigenvalues are the magnitudes of the principal stresses or strains.

Let's call the square matrix we are analyzing matrix $\boldsymbol{M}$.

$$\boldsymbol{M} = \begin{bmatrix} d & e \\ f & g \end{bmatrix}$$

We want to find all possible values for a variable we will call λ that satisfy the following ***characteristic equation***:

$$\det(\boldsymbol{M} - \lambda \boldsymbol{I}) = 0,$$

or using another common symbol for determinant,

$$\left| M - / I \right| = 0,$$

Matrix $\boldsymbol{I}$ in the preceding equations is the identity matrix

$$(\boldsymbol{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ if } \boldsymbol{M} \text{ is a 2x2 matrix})$$

and variable λ is an eigenvalue. Then

$$(\boldsymbol{M} - \lambda \boldsymbol{I}) = \begin{bmatrix} d & e \\ f & g \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} (d - \lambda) & e \\ f & (g - \lambda) \end{bmatrix}$$

The characteristic equation is

$$\begin{vmatrix} (d - \lambda) & e \\ f & (g - \lambda) \end{vmatrix} = 0$$

This determinant can be unpacked as the equation

$$((d - \lambda)(g - \lambda)) - (f\, e) = 0$$

or, simplifying,

$$\lambda^2 - ((d + g)\, \lambda\,) + ((d\, g\,) - (f\, e)) = 0.$$

The result is a quadratic equation. You might remember how to solve a quadratic equation from deep in your childhood. In case you don't, the general solution for the quadratic equation

$$(a\, x^2) + (b\, x) + c = 0$$

is

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

This expression, which is known as the *quadratic function*, yields two answers for x that satisfy the quadratic equation. In our case, the equivalent terms in the quadratic function are

$$a = 1$$

$$b = -(d + g),$$

noting that $[-(d + g)]^2 = (d + g)^2$, and

$$c = ((d\, g) - (f\, e))$$

Consequently, the eigenvalues are given by the quadratic function

$$\lambda = \frac{(d+g) \pm \sqrt{(d+g)^2 - 4(dg - fe)}}{2}$$

or, somewhat simplified,

$$\lambda = \frac{(d+g) \pm \sqrt{\left(4 * e * f\right) + \left(d - g\right)^2}}{2}.$$

We now have two values of λ that satisfy our quadratic equation, and these are the two eigenvalues of our 2x2 matrix. We will refer to the larger eigenvalue as λ_1, and the smaller eigenvalue is λ_2.

Now we need to find the eigenvectors that correspond to λ_1 and λ_2, respectively. Returning to our example using matrix ***M***, we have the following equation to solve to find the eigenvector associated with λ_1

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} d - \lambda_1 & e \\ f & g - \lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix},$$

which can be restated as

$$((d - \lambda_1)\, x_1) + (e\, y_1) = 0 \text{ and}$$

$$(f\, x_1) + ((g - \lambda_1)\, y_1) = 0.$$

The best we can do with these two equations and their two unknown values (x_1 and y_1) is to determine how one of the unknowns relates to the other. In other words, we can determine the slope of the vector. If we arbitrarily choose $x_1 = 1$, we can rearrange either or both of the previous equations to determine the value of y_1 when $x_1 = 1$:

$$y_1 = \frac{-\left(d - \lambda_1\right)}{e} = \frac{\left(\lambda_1 - d\right)}{e}$$

or

$$y_1 = \frac{-f}{\left(g - \lambda_1\right)}$$

An eigenvector that corresponds with eigenvalue λ_1 has the following coordinates:

$$\left\{1, \frac{\left(\lambda_1 - d\right)}{e}\right\}$$

The length or norm of that vector is

$$\sqrt{1^2 + \left(\frac{(\lambda_1 - d)}{e}\right)^2}$$

Recalling that a unit vector is a vector whose length is 1, we can find the unit vector associated with any vector of arbitrary length by dividing each component of the initial vector by that vector's length or norm. The unit eigenvector that corresponds to eigenvalue λ_1 is

$$\left\{\frac{1}{\sqrt{1^2 + \left(\frac{(\lambda_1 - d)}{e}\right)^2}}, \frac{\frac{(\lambda_1 - d)}{e}}{\sqrt{1^2 + \left(\frac{(\lambda_1 - d)}{e}\right)^2}}\right\}$$

We repeat the process to find the coordinates of the unit eigenvector that corresponds to eigenvalue λ_1.

Worked example.

Given the matrix ***M*** below, calculate the eigenvalues and the corresponding unit eigenvectors.

$$M = \begin{bmatrix} 6 & 3 \\ 3 & 4 \end{bmatrix}$$

Solution.

The eigenvalues are

$$\left(\frac{(6+4) + \sqrt{(4*3*3) + (6-4)^2}}{2}\right) = 8.16228$$

and

$$\left(\frac{(6+4) - \sqrt{(4*3*3) + (6-4)^2}}{2}\right) = 1.83772$$

We refer to the larger eigenvalue (8.16228) as λ_1, and the smaller (1.83772) is λ_2. An eigenvector associated with λ_1 is

$$\left\{1, \frac{(8.16228 - 6)}{3}\right\}.$$

The unit eigenvector associated with λ_1 is found by dividing the individual components of the previous vector by the vector's length

$$\left\{\frac{1}{\sqrt{1^2 + \left(\frac{(8.16228 - 6)}{3}\right)^2}}, \frac{\frac{(8.16228 - 6)}{3}}{\sqrt{1^2 + \left(\frac{(8.16228 - 6)}{3}\right)^2}}\right\} = \{0.811242, 0.58471\}$$

An eigenvector associated with λ_2 is

$$\left\{1, \frac{(1.83772 - 6)}{3}\right\}$$

and the corresponding unit eigenvector associated with λ_2 is

$$\left\{\frac{1}{\sqrt{1^2 + \left(\frac{(1.83772 - 6)}{3}\right)^2}}, \frac{\frac{(1.83772 - 6)}{3}}{\sqrt{1^2 + \left(\frac{(1.83772 - 6)}{3}\right)^2}}\right\} = \{0.58471, -0.811242\}$$

In the specific case in which we have a 2-D Lagrangian strain tensor

$$e_{ij} = \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}$$

the eigenvalues are given by

$$l = \frac{(e_{11} + e_{22}) \pm \sqrt{(4e_{12}e_{21}) + (e_{11} - e_{22})^2}}{2}.$$

The larger eigenvalue is λ_1, and the smaller eigenvalue is λ_2. An eigenvector corresponding to λ_1 is

$$\left\{1, \frac{\lambda_1 - e_{11}}{e_{12}}\right\},$$

and an eigenvector associated with λ_2 is

$$\left\{1, \frac{\lambda_2 - e_{11}}{e_{12}}\right\}.$$

The unit eigenvectors can then be determined by dividing each of the components of these vectors by their length or norm. The unit eigenvector associated with eigenvalue λ_1 is

$$\left\{\frac{1}{\sqrt{1+\left(\frac{(\lambda_1 - e_{11})}{e_{12}}\right)^2}}, \frac{\frac{(\lambda_1 - e_{11})}{e_{12}}}{\sqrt{1+\left(\frac{(\lambda_1 - e_{11})}{e_{12}}\right)^2}}\right\}$$

The unit eigenvector associated with eigenvalue λ_2 is

$$\left\{\frac{1}{\sqrt{1+\left(\frac{(\lambda_2 - e_{11})}{e_{12}}\right)^2}}, \frac{\frac{(\lambda_2 - e_{11})}{e_{12}}}{\sqrt{1+\left(\frac{(\lambda_2 - e_{11})}{e_{12}}\right)^2}}\right\}.$$

References

Ferguson, J., 1994, **Introduction to linear algebra in geology**: London, Chapman & Hall, 203 p., ISBN 0-412-49350-0.

Online resources

Eigenvalue calculator, accessed 25 June 2012 via http://www.wolframalpha.com/entities/calculators/eigenvalue_calculator/gv/9h/bi/

Weisstein, E.W., Eigenvalue: MathWorld—A Wolfram web resource, accessed 25 June 2012 via http://mathworld.wolfram.com/Eigenvalue.html

Weisstein, E.W., Eigenvector: MathWorld—A Wolfram web resource, accessed 25 June 2012 via http://mathworld.wolfram.com/Eigenvector.html

Weisstein, E.W., Quadratic equation: MathWorld—A Wolfram web resource, accessed 25 June 2012 via http://mathworld.wolfram.com/QuadraticEquation.html

Eigenvalue calculator, accessed 25 June 2012 via http://www.wolframalpha.com/entities/calculators/eigenvalue_calculator/gv/9h/bi/