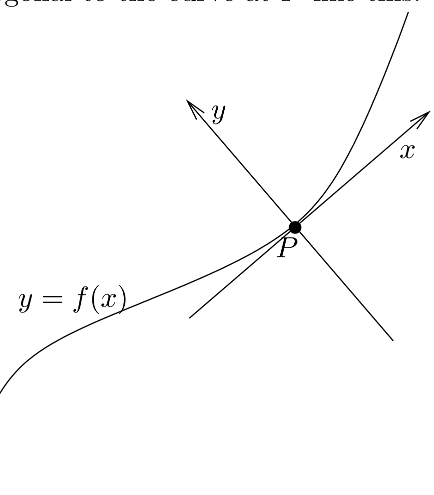
Mathematics 551 Homework, February 27, 2020

As a talked about in class on Wednesday one way we could have motivated the definition of the curvature of a curve, c, at a point, P, on would have been to choose a coordinate system with the x-axis tangent to the curve at P and the y-axis orthogonal to the curve at P like this:

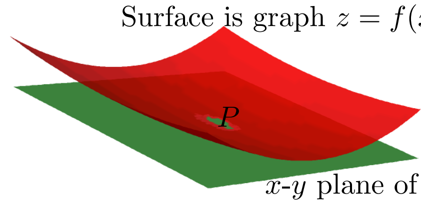


Then near P we can write the curve as the graph of a function $y = f(x)$. Because c goes through P, which is the origin of this coordinate system, $f(0) = 0$ and c is tangent to the x-axis at P which implies $f'(0) = 0$. Then we can define the curvature of c at P to be $f''(0)$. This agrees with our old definition as the formula for the curvature of a graph is

$$k(x) = \frac{f''(x)}{(1 + f'(x)^2)^{3/2}}$$

and using that $f'(0) = 0$ this gives $\kappa(0) = f''(0)$ at the origin.

For surfaces in $\mathbb{R}^3$ we can use the same idea. At a point, P of the surface choose coordinates so that the origin is at the point and x and y axis are tangent to the surface.



This time that the origin of the coordinate system is centered at P is and the x-y plane is tangent to the surface at P implies

$$f(0,0) = 0, \qquad \frac{\partial f}{\partial x}(0,0) = 0, \qquad \frac{\partial f}{\partial y}(0,0) = 0.$$

We can still use the second derivative of f at P as a measure of the curvature at P, but this time the second derivative is not just a scalar, but is the symmetric matrix

$$\begin{bmatrix} f_{xx}(0,0) & f_{yx}(0,0) \\ f_{xy}(0,0) & f_{yy}(0,0) \end{bmatrix}.$$

This homework introduces some linear algebra/matrix theory that will be useful for us in understanding the facts about symmetric matrices we will be using. On $\mathbb{R}^n$ (which for us will usually be $n = 2$ or $n = 3$ and for the discussion here I will use $n = 2$) write vectors as columns

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

The transpose of $\mathbf{x}$ is the row vector

$$\mathbf{x}^t = [x_1, x_2].$$

Then the inner product of two vectors can be written as

$$\mathbf{x} \cdot \mathbf{y} = [x_1, x_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1 y_1 + x_2 y_2.$$

It is more convenient to use a different notation for the inner product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y}.$$

For a square matrix

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

the transpose is

$$M^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

For any matrix (or vector) doing the transpose twice is the same as doing nothing:

$$(M^t)^t = M \qquad (\mathbf{x}^t)^t = \mathbf{x}.$$

Also taking transposes reverses the order of multiplication:

$$(M\mathbf{x})^t = \mathbf{x}^t M^t \qquad (MN)^t = N^t M^t.$$

Putting this all together gives

Proposition 1. *For any* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and any* $n \times n$ *matrix*

$$\langle \mathbf{x}, M\mathbf{y}\rangle = \langle M^t\mathbf{x}, \mathbf{y}\rangle \qquad \textit{and} \qquad \langle M\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, M^t\mathbf{y}\rangle.$$

Proof. Just compute

$$\langle \mathbf{x}, M\mathbf{y}\rangle = \mathbf{x}^t M\mathbf{y} = (M^t\mathbf{x})^t\mathbf{y} = \langle M^t\mathbf{x}, \mathbf{y}\rangle$$

and

$$\rangle M\mathbf{x}, \mathbf{y}\rangle = (M\mathbf{x})^t\mathbf{y} = \mathbf{x}M^t\mathbf{y} = \mathbf{x}(M^2\mathbf{y}) = \langle \mathbf{x}, M^t\mathbf{y}\rangle.$$

$\square$

That is taking the transpose of a matrix lets it jump over a comma in an inner product. A matrix, M, is ***symmetric*** if and only if $M^t = M$. In light of the last proposition this is the same as

$$\langle M\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, M\mathbf{y}\rangle.$$

Use a symmetric matrix can jump over commas without changing. While at first this dos not seem like much, we will see that it comes close to being a superpower.

Definition 2. A vector $\mathbf{v}$ is an ***eigenvector*** with ***eigenvalue*** λ (where λ is a scalar) of the matrix M if and only if $\mathbf{v} \neq \mathbf{0}$ and

$$M\mathbf{v} = \lambda\mathbf{v}.$$

$\square$

Here is the first of the powers of a symmetric matrix:

Proposition 3. *Let* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *be eigenvectors of the symmetric matrix* M *with distinct eigenvalues. That is*

$$M\mathbf{v}_1 = \lambda_1\mathbf{v}_1 \qquad M\mathbf{v}_2 = \lambda_2\mathbf{v}_1$$

with $\lambda_1 \neq \lambda_2$. *Then* $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *are orthogonal to each other.*

Proof. We need to show the inner product of the two vectors is zero: $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$. The proof does use anything other than the definitions and the comma hopping power of M.

$$\begin{aligned}
\lambda_1 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle &= \langle \lambda_1 \mathbf{v}_1, \mathbf{v}_2 \rangle \\
&= \langle M\mathbf{v}_1, \mathbf{v}_2 \rangle && (\text{as } \lambda_1 \mathbf{v}_1 = M\mathbf{v}_1) \\
&= \langle \mathbf{v}_1, M\mathbf{v}_2 \rangle && (\text{by comma hopping power}) \\
&= \langle \mathbf{v}_1, \lambda_1 \mathbf{v}_1 \rangle && (\text{as } M\mathbf{v}_2 = \lambda_2 \mathbf{v}_2) \\
&= \lambda_2 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle.
\end{aligned}$$

This can be rearranged to given

$$(\lambda_1 - \lambda_2)\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0.$$

Since $\lambda_1 \neq \lambda_2$, this implies $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$. □

The following is a special case of a more general result in linear algebra.

Proposition 4. *Let M be a 2×2 matrix. Then there is a nonzero vector $\mathbf{v}$ such that $M\mathbf{v} = \mathbf{0}$ if and only if $\det(M) = 0$.*

Proof. Let

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Then $M\mathbf{v} = \mathbf{0}$ if and only if

$$M\mathbf{v} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

First assume $M\mathbf{v} = 0$ with $\mathbf{v} \neq \mathbf{0}$. This implies

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned}$$

If $x \neq 0$ then multiply the first equation by d, and subtract b times the second equation (this is so that the y terms cancel) to get

$$(ad - bc)x = 0$$

which implies $\det M = (ad - bc) = 0$. If $x = 0$, then $y \neq 0$ (as $\mathbf{v} \neq \mathbf{0}$), then multiply the second equation by a and subtract c times the first equation to get

$$(ad - bc)y = 0$$

so that also in this case $\det M = ad - bc = 0$.

Conversely assume $\det M = ad - bc = 0$. If $M = 0$ is the matrix of all zeros, then $M\mathbf{v} = 0$ for all nonzero vectors $\mathbf{v}$. So assume that M has at least one nonzero element. Note

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} -b \\ a \end{bmatrix} = \begin{bmatrix} -ab + ab \\ -cb + da \end{bmatrix} = \begin{bmatrix} 0 \\ ad - bc \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d \\ -c \end{bmatrix} = \begin{bmatrix} ad-bc \\ cd-dc \end{bmatrix} = \begin{bmatrix} ad-bc \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$
and as not all of the entries of M are zero at least one of the vectors
$$\begin{bmatrix} -b \\ a \end{bmatrix}, \quad \text{or} \quad \begin{bmatrix} d \\ -c \end{bmatrix}$$
is not the zero vector and thus there is a nonzero vector $\mathbf{v}$ with $M\mathbf{v} = \mathbf{0}$. □

Proposition 5. *Let M be a 2×2 matrix. Then λ is an eigenvalue of M if and only if $\det(M - \lambda I) = 0$ (where I is the identity matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$).*

Proof. Note for a nonzero vector $\mathbf{v}$
$$\begin{aligned} M\mathbf{v} = \lambda\mathbf{v} &\iff M\mathbf{v} = \lambda I\mathbf{v} && (\text{as } I\mathbf{v} = \mathbf{v}) \\ &\iff M\mathbf{v} - \lambda I\mathbf{v} = \mathbf{0} \\ &\iff (M - \lambda I)\,\mathbf{v} = \mathbf{0}. \end{aligned}$$
And by Proposition 4 there is a nonzero vector $\mathbf{v}$ with $(M - \lambda I)\mathbf{v} = 0$ if and only if $\det(M - \lambda I) = 0$. □

The following is true for symmetric matrices of arbitrary size, but the proof requires more work.

Proposition 6. *If M is a 2×2 symmetric matrix, then all the solutions to $\det(M - \lambda I) = 0$ are real numbers.*

Proof. As M is symmetric it is of the form
$$M = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$
Then the equation we are interested in is
$$\begin{aligned} \det(M - \lambda I) &= \det\left(\begin{bmatrix} a-\lambda & b \\ b & c-\lambda \end{bmatrix}\right) \\ &= (a-\lambda)(b-\lambda) - b^2 \\ &= \lambda^2 - (a+c)\lambda + ac - b^2 \\ &= 0. \end{aligned}$$

Problem 1. Show that the solutions to this equation are
$$\lambda = \frac{(a+c) \pm \sqrt{(a-c)^2 + 4b^2}}{2}$$
□

As $(a-c)^2 + 4b^2 \geq 0$ the square root $\sqrt{(a-c)^2 + 4b^2}$ is a real number and so the roots are real. □

Theorem 7 (Principle Axis Theorem)**.** *Let M be a 2×2 symmetric matrix. Then there are orthonormal vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ (that is each has length 1 and they are orthogonal to each other) and real number λ_1 and λ_2 so that*
$$M\mathbf{u}_1 = \lambda_1\mathbf{u}_1 \qquad \textit{and} \qquad M\mathbf{u}_2 = \lambda_2\mathbf{u}_2.$$

Problem 2. Prove this. *Hint:* Let λ_1 and λ_2 be the solutions to $\det(M - \lambda I) = 0$. By Proposition 6 these are real numbers. First assume that $\lambda_1 \neq \lambda_2$. By Proposition 5 there are nonzero vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ with

$$M\mathbf{v}_1 = \lambda_1 \mathbf{v}_1 \qquad \text{and} \qquad M\mathbf{v}_2 = \lambda_2 \mathbf{v}_2.$$

Use Proposition 3 to show

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 \qquad \text{and} \qquad \mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2$$

do the trick.

If $\lambda_1 = \lambda_2$ let $\lambda = \lambda_1$ and show that $M = \lambda I$ and therefore any pair of orthonormal vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ will do the trick (note if $\lambda_1 = \lambda_2$, then by the formula of Problem 1 $(a-c)^2 + b^2 = 0$).

For any θ let

$$P(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

be the rotation by θ.

Problem 3. Show

$$P(\alpha)P(\beta) = P(\alpha+\beta). \qquad \square$$

It very easy to find the inverse of a rotation:

Proposition 8. *Let $P = P(\theta)$ be a rotation. Then*

$$P^t P = I$$

and thus the transpose, P^t, is the inverse of P.

Problem 4. Prove this. $\square$

Theorem 9. *Let M be a symmetric matrix and let $\mathbf{e}_1$ and $\mathbf{e}_2$ be the standard basis of $\mathbb{R}^2$. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ orthogonal basis of $\mathbb{R}^2$ given by Theorem 7. Let $P = P(\theta)$ be a rotation so that*

$$P\mathbf{e}_1 = \mathbf{u}_1, \qquad P\mathbf{e}_2 = \mathbf{u}_2$$

and let

$$N = P^t M P = P^{-1} M P.$$

Then

$$N\mathbf{e}_1 = \lambda_1 \mathbf{e}_1, \qquad N\mathbf{e}_2 = \lambda \mathbf{e}_2.$$

(In linear algebra terminology this shows that M is similar to a diagonal matrix and shows that the matrix of N in the standard basis is the diagonal matrix $\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$).

Problem 5. Prove this. $\square$

Recall that the ***trace*** of a matrix is the sum of the diagonal entries, that is

$$\operatorname{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d.$$

Proposition 10. *Let A and B are square matrices then*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA) \qquad \textit{and} \qquad \det(AB) = \det(BA).$$

This implies that if $N = P^{-1}MP$, then

$$\operatorname{tr}(N) = \operatorname{tr}(M) \qquad \textit{and} \qquad \det(N) = \det(M).$$

Proof. We will assume that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ and $\det(AB) = \det(BA)$ are know results. Then

$$\operatorname{tr}(N) = \operatorname{tr}(P^{-1}MP) = \operatorname{tr}((P^{-1}M)P) = \operatorname{tr}(P(P^{-1}M)) = \operatorname{tr}(IM) = \operatorname{tr}(M).$$

A similar calculation shows $\det(N) = \det(M)$. $\square$