Analyzing the Hessian

- Premise
- Determinants
- Eigenvalues
- Meaning

The Problem

In 1-variable calculus, you can just look at the second derivative at a point and tell what is happening with the concavity of a function: positive implies concave up, negative implies concave down.

But because the Hessian (which is equivalent to the second derivative) is a matrix of values rather than a single value, there is extra work to be done.

This lesson forms the background you will need to do that work.

Finding a Determinant

Given a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant, symbolized $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, is equal to a·d - b·c. So, the determinant of

$$\begin{bmatrix} 3 & 4 \\ -1 & 2 \end{bmatrix} \text{ is... } \quad 6 \quad - \quad -4 = 10$$

The determinant has applications in many fields. For us, it's just a useful concept.

Determinants of larger matrices are possible to find, but more difficult and beyond the scope of this class.

Practice Problem 1

Find the determinant. Check your work using `det(A)` in Julia.

a. $\begin{vmatrix} 3 & 1 \\ -2 & 0 \end{vmatrix}$

b. $\begin{vmatrix} 4 & 1 \\ 1 & 5 \end{vmatrix}$

c. $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$

Eigenvectors and Eigenvalues

One of the biggest applications of matrices is in performing geometric transformations like rotation, translation, reflection, and dilation.

In Julia, type in the vector `X = [3; -1]`

Then, multiply `[2 0; 0 2]*X`

You should get [6; -2], which is a multiplication of X by a factor of 2, in other words a dilation.

Next, try $\begin{bmatrix} \mathtt{cos(\frac{pi}{6})} & \mathtt{-sin(\frac{pi}{6})} \\ \mathtt{sin(\frac{pi}{6})} & \mathtt{cos(\frac{pi}{6})} \end{bmatrix}\mathtt{*X}$

Although this one isn't immediately clear, you have accomplished a rotation of vector X by $\pi/6$ radians.

Eigenvectors and Eigenvalues

In the last slide, we were looking at a constant X and a changing A, but you can also get interesting results for a constant A and a changing X.

For example, the matrix $\begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix}$ doesn’t look very special, and it doesn’t do anything special for most values of X.

But if you multiply it by $\begin{bmatrix} 3 \\ 5 \end{bmatrix}$, you get $\begin{bmatrix} 21 \\ 35 \end{bmatrix}$, which is a scalar multiplication by 7.

Eigenvectors and Eigenvalues

When a random matrix A acts as a scalar multiplier on a vector X, then that vector is called an *eigenvector* of X.

The value of the multiplier is known as an *eigenvalue*.

For the purpose of analyzing Hessians, the eigenvectors are not important, but the eigenvalues are.

Finding Eigenvalues

The simplest way to find eigenvalues is to open Julia and type in:

```
eig(A)
```

This will give you the eigenvalue(s) of A as well as a matrix composed of the associated eigenvectors.

However, it's also useful to know how to do it by hand.

Finding Eigenvalues

To find eigenvalues by hand, you will be solving this equation…

determinant symbol

$$\left| \underbrace{\begin{bmatrix} a & b \\ c & d \end{bmatrix}}_{\text{original matrix}} - \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} \right| = 0$$

variable matrix, will solve for x

…which turns into the following determinant:

$$\begin{vmatrix} a - x & b \\ c & d - x \end{vmatrix} = 0$$

Finding Eigenvalues

So, if you were trying to find the eigenvalues for the matrix $\begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix}$, you would need to solve the determinant $\begin{vmatrix} 2-x & 3 \\ 5 & 4-x \end{vmatrix} = 0$.

Cross-multiplying, you would get

$(2 - x)(4 - x) - 15 = 0$

$8 - 6x + x^2 - 15 = 0$

$x^2 - 6x - 7 = 0$

$(x - 7)(x + 1) = 0$ so $\boxed{x = 7 \text{ or } -1.}$

eigenvalues!

Practice Problem 2

Find the eigenvalues of the following matrices by hand, then check using Julia:

a. $\begin{bmatrix} 3 & 8 \\ 4 & -1 \end{bmatrix}$

b. $\begin{bmatrix} 2 & 6 \\ -1 & 3 \end{bmatrix}$

Find the eigenvalues using Julia:

c. $\begin{bmatrix} 2 & 1 & -4 \\ -2 & 3 & -1 \\ 0 & 1 & -2 \end{bmatrix}$

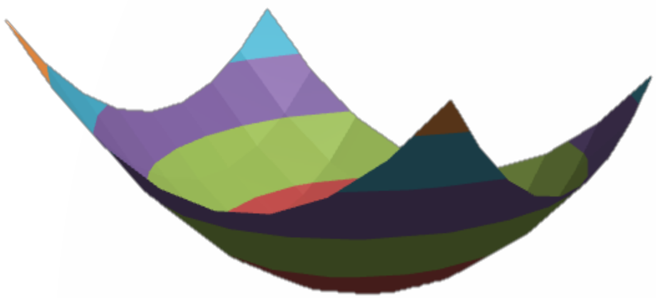
Meaning of Eigenvalues

Because the Hessian of an equation is a square matrix, its eigenvalues can be found (by hand or with computers – we'll be using computers from here on out).

Because Hessians are also symmetric (the original and the transpose are the same), they have a special property that their eigenvalues will always be real numbers.

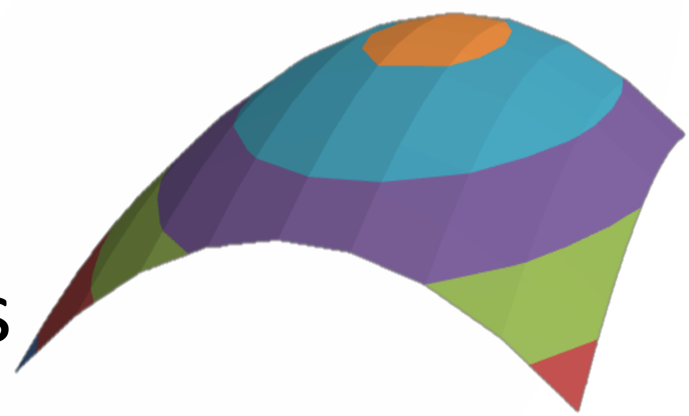
So the only thing of concern is whether the eigenvalues are *positive* or *negative*.

Meaning of Eigenvalues



If the Hessian at a given point has all positive eigenvalues, it is said to be a positive-definite matrix. This is the multivariable equivalent of “concave up”.

If all of the eigenvalues are negative, it is said to be a negative-definite matrix. This is like “concave down”.



Meaning of Eigenvalues

If either eigenvalue is 0, then you will need more information (possibly a graph or table) to see what is going on.

And, if the eigenvalues are mixed (one positive, one negative), you have a saddle point:

Here, the graph is concave up in one direction and concave down in the other.

Practice Problem 3

Use Julia to find the eigenvalues of the given Hessian at the given point. Tell whether the function at the point is concave up, concave down, or at a saddle point, or whether the evidence is inconclusive.

a. $\begin{bmatrix} 12x^2 & -1 \\ -1 & 2 \end{bmatrix}$ at (3, 1)

b. $\begin{bmatrix} 6x & 0 \\ 0 & 6y \end{bmatrix}$ at (-1, -2)

c. $\begin{bmatrix} -2y^2 & -4xy \\ -4xy & -2x^2 \end{bmatrix}$ at (1, -1) and (1, 0)

Practice Problem 4

Determine the concavity of

$$f(x, y) = x^3 + 2y^3 - xy$$

at the following points:

a) (0, 0)
b) (3, 3)
c) (3, -3)
d) (-3, 3)
e) (-3, -3)

Practice Problem 5

For $f(x, y) = 4x + 2y - x^2 - 3y^2$

a) Find the gradient. Use that to find a critical point (x, y) that makes the gradient 0.

b) Use the eigenvalues of the Hessian at that point to determine whether the critical point in a) is a maximum, minimum, or neither.

Practice Problem 6

For $f(x, y) = x^4 + y^2 - xy$,

a) Find the critical point(s)

b) Test the critical point(s) to see if they are maxima or minima.