Math HW5

Matthew Vanaman

03-27-2019

5.2, 5.3

4.8

(a)

Work

First Derivative:

The only layer here is in the exponent of e; everything else is constant.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

$$A = -\frac{x^2}{2}, \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{A}$$

$$A' = \frac{d}{dx}\Big(-\frac{x^2}{2}\Big) = -\frac{1}{2}\Big(2x^{2-1}\Big) = -\frac{2x^{2-1}}{2} = -x.$$

$$\frac{1}{\sqrt{2\pi}} e^{A}\Big(-x\Big) = \frac{-x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Second Derivative

Because the first derivative was multiplicative, we need the product to get the second derivative.
Product rule: $f''(x) = g'(x)f(x) + g(x)f'(x)$.
$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.
$f'(x) = \frac{-x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.
$g(x) = -x$.
$g'(x) = \frac{d}{dx}(-x) = -1$.

$$\begin{aligned} f''(x) &= -1\Bigg(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\Bigg) - x\Bigg(\frac{-x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\Bigg) \\ &= -\Bigg(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\Bigg) + \Bigg(\frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\Bigg) \\ &= \Bigg(\frac{x^2-1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\Bigg). \end{aligned}$$

Plug in 0 for each version

$$f(0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{0^2}{2}} = \frac{1}{\sqrt{2\pi}}.$$

$$f'(0) = \frac{0^2}{\sqrt{2\pi}} e^{-\frac{0^2}{2}} = 0.$$

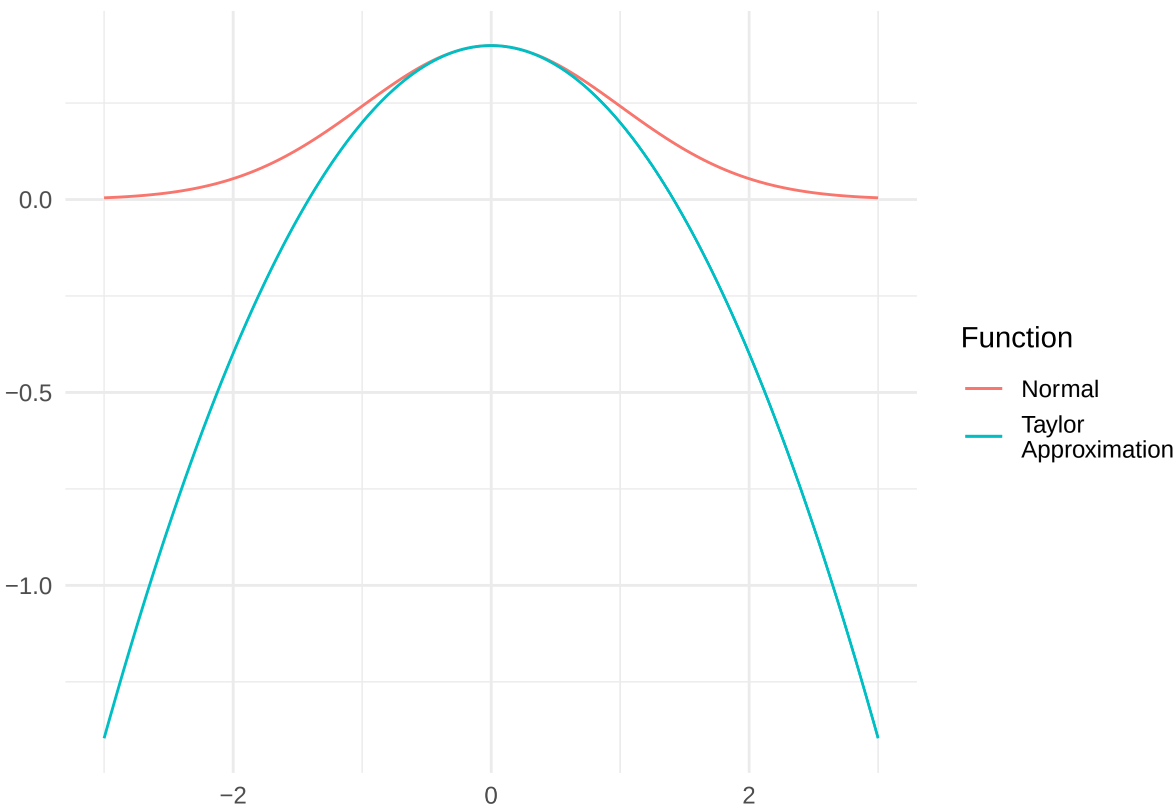
$$f''(0) = \frac{0^2 - 1}{\sqrt{2\pi}} e^{-\frac{0^2}{2}} = \frac{-1}{\sqrt{2\pi}}.$$

Taylor approximation - just plug in what we found:

$$f(x) \approx \frac{1}{\sqrt{2\pi}} + 0x + \frac{-1/\sqrt{2\pi}}{2} x^2.$$

$$f''(x) \approx \frac{-1}{\sqrt{8\pi}} x^2 + \frac{1}{\sqrt{2\pi}}.$$

(b)



Taylor series seems most accurate between about -1 and 1.

4.9

(a)

Start with

$$\ln\left(\prod_{i=1}^{N} p^{y_i}(1-p)^{1-y_i}\right),$$

$$\sum_{i=1}^{N}\ln\left(p^{y_i}(1-p)^{1-y_i}\right),$$

$$\sum_{i=1}^{N}\left(\ln\left(p^{y_i}\right)+\ln\left((1-p)^{1-y_i}\right)\right),$$

$$\sum_{i=1}^{N}\left((y_i)\ln\left(p\right)+\ln\left((1-p)^{1-y_i}\right)\right),$$

$$\sum_{i=1}^{N}\left((y_i)\ln\left(p\right)+(1-y_i)\ln(1-p)\right),$$

$$\sum_{i=1}^{N}(y_i)\ln\left(p\right)+\sum_{i=1}^{N}(1-y_i)\ln(1-p),$$

$$\ln\left(p\right)\sum_{i=1}^{N}(y_i)+\ln(1-p)\sum_{i=1}^{N}(1-y_i),$$

$$\ln\left(p\right)\sum_{i=1}^{N}(y_i)+\ln(1-p)\left(\sum_{i=1}^{N}1\right)-\left(\sum_{i=1}^{N}y_i\right),$$

$$\ln\left(p\right)\sum_{i=1}^{N}(y_i)+\ln(1-p)\left(N-\sum_{i=1}^{N}y_i\right).$$

(b)

$$\frac{d}{dp}\left(\ln\left(p\right)\sum_{i=1}^{N}(y_i)+\ln(1-p)\left(N-\sum_{i=1}^{N}y_i\right)\right),$$

$$\frac{d}{dp}\left(\ln\left(p\right)\sum_{i=1}^{N}(y_i)\right)+\frac{d}{dp}\left(\ln(1-p)\left(N-\sum_{i=1}^{N}y_i\right)\right),$$

$$\frac{d}{dp}\left(\ln\left(p\right)\right)\frac{d}{dp}\left(\sum_{i=1}^{N}(y_i)\right)+\frac{d}{dp}\left(\ln(1-p)\right)\frac{d}{dp}\left(N-\sum_{i=1}^{N}y_i\right),$$

$$\frac{d}{dp}\left(\ln\left(p\right)\right)\frac{d}{dp}\left(\sum_{i=1}^{N}(y_i)\right)+\frac{d}{dp}\left(\ln(1-p)\right)\frac{d}{dp}\left(N-\sum_{i=1}^{N}y_i\right),$$

$$\frac{1}{p}\left(\sum_{i=1}^{N}(y_i)\right)+\frac{d}{dp}\left(\ln(1-p)\right)\left(N-\sum_{i=1}^{N}y_i\right).$$

Gotta use the chain rule:

$$\frac{d}{dp}\Bigg(\ln(1-p)\Bigg) = \left(\frac{1}{1-p}\right)\frac{d}{dp}(1-p) = \frac{\frac{d}{dp}(1) - \frac{d}{dp}(p)}{1-p} = \frac{0-1}{1-p} = -\frac{-1}{1-p}.$$

$$\frac{1}{p}\left(\sum_{i=1}^{N}(y_i)\right) + \frac{-1}{1-p}\left(N - \sum_{i=1}^{N} y_i\right),$$

$$\frac{1}{p}\left(\frac{\sum_{i=1}^{N}(y_i)}{1}\right) + \frac{-1}{1-p}\left(\frac{N - \sum_{i=1}^{N} y_i}{1}\right),$$

$$\frac{\sum_{i=1}^{N}(y_i)}{p} - \frac{N - \sum_{i=1}^{N} y_i}{1-p}.$$

(c)

$$0 = \frac{\sum_{i=1}^{N}(y_i)}{p} - \frac{N - \sum_{i=1}^{N} y_i}{1-p},$$

$$\frac{\sum_{i=1}^{N}(y_i)}{p} = \frac{N - \sum_{i=1}^{N} y_i}{1-p},$$

$$\frac{\sum_{i=1}^{N}(y_i)}{p} = \frac{N - \sum_{i=1}^{N} y_i}{1-p},$$

$$\sum_{i=1}^{N} y_i(1-p) = p\left(N - \sum_{i=1}^{N} y_i\right),$$

$$\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_i(p) = (p)N - (p)\sum_{i=1}^{N} y_i,$$

$$\sum_{i=1}^{N} y_i = (p)N,$$

$$\frac{\sum_{i=1}^{N} y_i}{N} = p.$$

In the maximum likelihood framework, the proportion observed in the data tracks the probability of in the population, assuming assumptions hold. In this case, we have a sample where we add up all of the y values, the divide by the number of values. This gives us the sample proportion (or mean), which is an estimate of the probability.

(d)

If $p = \frac{\sum_{i=1}^{N} y_i}{N}$, and we're given a function that gives us p, then all we have to do is substitute the function for p, then substitute x_i into that function.

$$p = \frac{1}{1 + e^{-(0.2+0.5x_i)}}.$$

Very Conservative:

$$p = \frac{1}{1 + e^{-(0.2+0.53)}} = 0.85.$$

```r
## calculations
# get euler's number
exp(1)
```

```
[1] 2.718282
```

```r
# get p for Very conservative:
1 / (1 + exp(1)^-(0.2 + (0.5*3)))
```

```
[1] 0.8455347
```

Moderate:

$$p = \frac{1}{1 + e^{-(0.2+0.5(0))}} = 0.55.$$

```r
## calculations
# get p for Moderate:
1 / (1 + exp(1)^-(0.2 + (0.5*0)))
```

```
[1] 0.549834
```

Very Liberal:

$$p = \frac{1}{1 + e^{-(0.2+0.5(-3))}} = 0.21.$$

```r
## calculations
# get p for Very Liberal:
1 / (1 + exp(1)^-(0.2 + (0.5*(-3))))
```

```
[1] 0.214165
```

(e)

Predicted probabilities (?):

$$p' = \frac{d}{dx}\left(\frac{1}{1 + e^{-(0.2+0.5x_i)}}\right).$$

Need chain rule...

$$p = \frac{1}{A}, \quad A = 1 + e^B, \quad B = -(0.2 + 0.5x_i).$$

$$p' = \frac{d}{dp}\left(\frac{1}{A}\right) = -\frac{1}{A^2}.$$

$$A' = \frac{d}{dA}(1 + e^B) = \frac{d}{dA}(1) + \frac{d}{dA}(e^B) = 0 + e^B = e^B.$$

$$B' = \frac{d}{dB}[-(0.2 + 0.5x_i)] = \frac{d}{dB}(-0.2) + (-0.5)\frac{d}{dB}(x_i) = 0 - 0.5 \times 1 = -0.5.$$

$$\frac{(-1)(e^B)(-0.5)}{A^2} = \frac{(0.5)e^B}{A^2},$$

$$\frac{(0.5)e^B}{A^2} = \frac{(0.5)e^B}{(1 + e^B)^2},$$

$$\frac{(0.5)e^B}{(1 + e^B)^2} = \frac{0.5e^{-(0.2+0.5x_i)}}{(1 + e^{-(0.2+0.5x_i)})^2}.$$

Plug in zero:

$$\frac{0.5e^{-(0.2+0.5(0))}}{(1+e^{-(0.2+0.5(0))})^2} = 0.124.$$

```r
# calculation
0.5 * exp(1)^-(0.2+(0.5 * 0)) / (1 + exp(1)^-(0.2+(0.5 * 0)))^2
```

```
[1] 0.1237583
```

(f)

This funciton gives you the instantaneous rate of change for a given x value. This means that for a moderate voter (x = 0), the instantaneous rate of change in probability of voting for the incumbant is 0.124. In other words, it is the slope of the line at a given point of x.

5.1

(a)

Take derivative:

$$f(x) = 3x^4 - 4x^3 - 36x^2,$$

$$f'(x) = (4)3x^{4-1} - (3)4x^{3-1} - (2)36x^{2-1} = 12x^3 - 12x^2 - 72x.$$

$$\begin{aligned} 0 &= 12x^3 - 12x^2 - 72x \\ &= 12(x^3 - x^2 - 6x) \\ &= 12x(x^2 - x - 6) \\ &= 12x(x-3)(x+2). \end{aligned}$$

Critical points are when x equals 3, -2, or 0. Using second derivitive test, use the critical points rules:

$$f''(x) = (3)12x^{3-1} - (2)12x^{2-1} - 72x^{1-1} = 36x^2 - 24x - 72.$$

For x = -2:

$$36(-2)^2 - 24(-2) - 72 = 120$$

```r
# calculation
36 * (-2)^2 - 24 * (-2) - 72
```

```
[1] 120
```

For x = 3

$$36(3)^2 - 24(3) - 72 = 180$$

```r
36*3^2 - 24 * 3 - 72
```

```
[1] 180
```

For x = 0:

$$36(0)^2 - 24(0) - 72 = -72$$

```r
36*0^2 - 24 * 0 - 72
```

```
[1] -72
```

According to the second derivative test, a critical point is a local maximum if $f''(x) < 0$, a local minimum if $f''(x) > 0$, and is a saffle point when $f''(x) = 0$. This means critical points x = 3 and -2 are local minimums while x = 0 is a local maximum.
to find global max and min, first plug in the boundary points into the original function.
Lower boundary:

$$f(-4) = 3(-4)^4 - 4(-4)^3 - 36(-4)^2 = 448$$

```
# calculation
3*(-4)^4 - 4*(-4)^3 - 36*(-4)^2
```

```
[1] 448
```

Upper boundary:

$$f(4) = 3(4)^4 - 4(4)^3 - 36(4)^2 = -64$$

```
# calculation
3*(4)^4 - 4*(4)^3 - 36*(4)^2
```

```
[1] -64
```

Finally, compare these to the local minimum and maximum points:
local maximum:

$$f(0) = 3(0)^4 - 4(0)^3 - 36(0)^2 = 0.$$

```
3*(0)^4 - 4*(0)^3 - 36*(0)^2
```

```
[1] 0
```

Local minima:

$$f(3) = 3(3)^4 - 4(3)^3 - 36(3)^2 = -189.$$
$$f(-2) = 3(-2)^4 - 4(-2)^3 - 36(-2)^2 = -64.$$

```
3*(3)^4 - 4*(3)^3 - 36*(3)^2
```

```
[1] -189
```

```
3*(-2)^4 - 4*(-2)^3 - 36*(-2)^2
```

```
[1] -64
```

Output of function is lowest at x = (3), therefore this is the global minimum. Output of function is highest at x = -4, therefore this is global maximum.

(b)

*lower boundary approaches 0.

$$\begin{aligned}
g'(x) &= \frac{d}{dx}(x\ln(x) - x), \\
&= \frac{d}{dx}\big(x\ln(x)\big) - \frac{d}{dx}(x), \\
&= (x)\frac{d}{dx}\ln(x) - 1, \\
&= (x)\frac{1}{x} + \ln(x) - 1, \\
&= 1 + \ln(x) - 1,
\end{aligned}$$

$$= \ln(x).$$

Set derivative equal to zero:

$$0 = \ln(x) \text{ when } x = 1.$$

Is critical point positive or negative?

$$g''(x) = \frac{x}{dx}\Big(\ln(x)\Big) = \frac{1}{x},$$

$$g''(x) = \frac{x}{dx}\Big(\ln(1)\Big) = \frac{1}{1} = 1.$$

Positve, therefore local minimum. Now compare critical point $x = 1$ with boundary point in original function:

$$f(1) = (1)\ln(1) - 1 = -1.$$

```
1 * log(1) - 1
```

```
[1] -1
```

$$f(3) = (3)\ln(3) - 3 = 2.30$$

```
(3) * log(3) - 3
```

```
[1] 0.2958369
```

Critical point $x = 1$ is global minimum (-1) and $x = 3$ is global maximum (2.30).

5.2

(a)

Take the derivative:

$$f'(x) = 3x^2 - 15x + 12,$$
$$= 3(x^2 - 5 + 4),$$
$$= 3(x - 1)(x - 4).$$

We have critical points 1 and 4.
First derivative test: left of x is positive:

$$f'(0) = 3(0)^2 - 15(0) + 12 = 12$$

```
3*(0)^2 - 15*(0) + 12
```

```
[1] 12
```

Right of x is negative:

$$= 3(2)^2 - 15(2) + 12 = -6$$

```
3*(2)^2 - 15*(2) + 12
```

```
[1] -6
```

Critical point x = 1 is a local maximum.
Critical point 4, to the left is negative:

$$f'(3) = 3(3)^2 - 15(3) + 12 = -6$$

```r
3*(3)^2 - 15*(3) + 12
```

```
[1] -6
```

And right is positive:

$$f'(5) = 3(5)^2 - 15(5) + 12 = 12$$

```r
3*(5)^2 - 15*(5) + 12
```

```
[1] 12
```

Critical point x = 4 is a local minimum.
Compare critical points against boundary points:

$$f(x) = x^3 - \frac{15}{2}x^2 + 12x + 8$$

At lower boundary:

$$f(0) = 0^3 - \frac{15}{2}0^2 + 12(0) + 8 = 8.$$

```r
0^3 - {15}/{2} * 0^2 + 12*(0) + 8
```

```
[1] 8
```

At upper boundary:

$$f(6) = 6^3 - \frac{15}{2}6^2 + 12(6) + 8 = 26.$$

```r
6^3 - {15}/{2} * 6^2 + 12*(6) + 8
```

```
[1] 26
```

At critical point x = 1:

$$f(1) = 1^3 - \frac{15}{2}1^2 + 12(1) + 8 = 13.5.$$

```r
1^3 - {15}/{2} * 1^2 + 12*(1) + 8
```

```
[1] 13.5
```

At critical point x = 4:

$$f(4) = 4^3 - \frac{15}{2}4^2 + 12(4) + 8 = 0.$$

```r
4^3 - {15}/{2} * 4^2 + 12*(4) + 8
```

```
[1] 0
```

Critical points x = 4 and x = 6 are the locations of the global minimum and maximum, respectively.

(b)

Starting at 2:

Iteration	x	fp(x)	fpp(x)	fNR
0	2.0000000	-6.0000000	-3.000000	0.0000000
1	0.0000000	12.0000000	-15.000000	0.8000000
2	0.8000000	1.9200000	-10.200000	0.9882353
3	0.9882353	0.1062975	-9.070588	0.9999542
4	0.9999542	0.0004122	-9.000275	1.0000000
5	1.0000000	0.0000000	-9.000000	1.0000000

Converges after 5, at root 1.

Starting at 5:

Iteration	x	fp(x)	fpp(x)	fNR
0	5.000000	12.0000000	15.000000	4.200000
1	4.200000	1.9200000	10.200000	4.011765
2	4.011765	0.1063002	9.070590	4.000046
3	4.000046	0.0004140	9.000276	4.000000
4	4.000000	0.0000000	9.000000	4.000000

Converges after 4, at root 4.

(c)

One problem is that there could be more than one root. If you don't know this, you may have missed out on another convergence point and come to the wrong conclusion that the convergence point was the only root, or the best root (on substantive grounds). Unfortuntaly there's no way to tell unless you try out different starting values.

Another problem is that the convergence point is itself limited in the information it gives you. If there are multiple convergence points, that could speak to the type of curve that function produces e.g. a cubic function has a local maximum and global maximum, but you don't necessarily know whether the convergence point arrived at the local or global maximum. You could potentially misunderstand what the function is telling you.