Chapter 5: Optimization

1. (a) The notation $x \in [-4, 4]$ means that x is bounded, and that both boundary points are included in the domain. That means we have to compare the value of the function at the critical points to the value of the function at the boundary points to find the global maximum and global minimum. First we find the critical points by taking the derivative, setting it equal to 0, and solving for x. The derivative is

$$f'(x) = 12x^3 - 12x^2 - 72x.$$

If we set the derivative equal to 0, we can solve for x through factoring:

$$12x^3 - 12x^2 - 72x = 0,$$

$$12(x^3 - x^2 - 6x) = 0,$$
$$12x(x^2 - x - 6) = 0,$$
$$12x(x^2 - x - 6) = 0,$$
$$12x(x - 3)(x + 2) = 0.$$

So the critical points are $x = 0$, $x = 3$, and $x = -2$. Next, we can check whether each critical point c describes a local maximum, a local minimum, or a saddle point using the second derivative test. The second derivative of the function is

$$f''(x) = 36x^2 - 24x - 72.$$

c	$f''(c)$	Result
0	$36(0)^2 - 24(0) - 72 = -72$	Local max
3	$36(3)^2 - 24(3) - 72 = 180$	Local min
-2	$36(-2)^2 - 24(-2) - 72 = 120$	Local min

Finally, we compare the local maximum to the boundary points to find the global maximum, and we compare the local minimums to the boundary points to find the global minimum. At the boundary points, the function is

$$f(-4) = 3(-4)^4 - 4(-4)^3 - 36(-4)^2 = 448, \qquad f(4) = 3(4)^4 - 4(4)^3 - 36(4)^2 = -64,$$

for the local maximum the function is

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

and at the local minimums the function is

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

So the global maximum occurs at the boundary point $x = -4$, and the global minimum occurs at $x = 3$.

(b) Since the domain is bounded and since the upper bound 3 is included in the domain (the parenthesis around 0 means that it is not included in the domain) we have to compare the value of the function at $x = 3$ to the value of the function at the critical points. To find the critical points, we take the derivative. First note that the derivative can be broken up across subtraction:

$$g'(x) = \frac{d}{dx}\Big(x\ln(x)\Big) - \frac{d}{dx}(x)$$

$$= \frac{d}{dx}\Big(x\ln(x)\Big) - 1.$$

The product rule applies

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

Next we set the derivative equal to 0. We know that any logarithm of 1 is equal to 0, so the critical point is $x = 1$. To test whether this point is a local max or a local min, we find the second derivative,

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

and plug the critical point in:

$$g'(1) = \frac{1}{1} = 1.$$

Since the second derivative is positive at the critical point, $x = 1$ describes a local minimum. Finally we compare the value of the function at the local min to the boundary point:

$$f(1) = (1)\ln(1) - 1 = 1(0) - 1 = -1,$$

$$f(3) = (3)\ln(3) - 1 = 2.3.$$

So $x = 1$ is the location of the global minimum, and $x = 3$ is the location of the global maximum.

2. (a) First we find the critical points by taking the derivative and setting it equal to 0:

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

We can factor out a 3:

$$3(x^2 - 5x + 4) = 0.$$

Two numbers that add to -5 and multiply to 4 are -1 and -4, so the derivative factors to

$$3(x-1)(x-4) = 0.$$

That implies that the critical points are $x = 1$ and $x = 4$, both of which exist in the stated domain. Next we check whether each one is a local minimum, local maximum, or a saddle point. The second derivative is a pretty simple function in this case, so we use the second derivative test by finding the second derivative, plugging in the critical points, and observing whether the second derivative is positive or negative:

$$f''(x) = 6x - 15.$$

The second derivative at $x = 1$ is

$$f''(1) = 6(1) - 15 = -9,$$

so $x = 1$ represents a local maximum. The second derivative at $x = 5$ is

$$f''(4) = 6(4) - 15 = 9,$$

so $x = 4$ is a local minimum. Finally, we compare the value of the function at the critical points to the value of the function at the boundary points:

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

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

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

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

So $x = 5$ is the global minimum, but $x = 6$ is the global maximum.

(b) The Newton-Raphson algorithm, beginning at 2, yields the following iterations:

Iteration	x	$f'(x)$	$f''(x)$	$x - \frac{f'(x)}{f''(x)}$
0	2	-6	-3	0
1	0	12	-15	0.8
2	0.8	1.92	-10.2	0.988235294
3	0.988235294	0.106297578	-9.070588235	0.999954223
4	0.999954223	0.000412	-9.000274662	0.999999999
5	0.999999999	6.28643E-09	-9.000000004	1
6	1	0	-9	1

The Newton-Raphson algorithm, beginning at 5, yields the following iterations:

Iteration	x	$f'(x)$	$f''(x)$	$x - \frac{f'(x)}{f''(x)}$
0	5	12	15	4.2
1	4.2	1.92	10.2	4.011764706
2	4.011764706	0.106297578	9.070588235	4.000045777
3	4.000045777	0.000412	9.000274662	4.000000001
4	4.000000001	6.28643E-09	9.000000004	4
5	4	0	9	4

(c) i. The NR-algorithm might stop at a minimum or a maximum, and does not tell you which one it arrives at.

ii. The NR-algorithm might stop at a local extreme point that is not a global extreme point. If the global max or min is on the boundary, then the NR-algorithm will never return that point.

3. To find the global maximum of $U(t)$, we first have to find the critical points. Technically, t is bounded below by 0, so we have to compare the value of the function at the critical points to the value at 0. To find the critical points, we first take the derivative:

$$U'(t) = \frac{d}{dt}\left(30\ln(t+1) - \frac{t^2}{10}\right)$$

$$= 30\frac{d}{dt}\left(\ln(t+1)\right) - \frac{\frac{d}{dt}(t^2)}{10}$$

$$= \frac{30}{t+1} - \frac{t}{5}.$$

Next we set the derivative equal to 0 and we solve for t:

$$\frac{30}{t+1} - \frac{t}{5} = 0,$$

$$\frac{30}{t+1} = \frac{t}{5},$$

$$t(t+1) = 150,$$

$$t^2 + t - 150 = 0.$$

This quadratic equation does not factor neatly, but we can use the quadratic formula to calculate the solutions. The quadratic formula says that for an equation of the form $Ax^2 + Bx + C = 0$, x is

$$x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

Let $A = 1$, $B = 1$, and $C = -150$. Then t is

$$t = \frac{-1 \pm \sqrt{1^2 - 4(1)(-150)}}{2(1)}$$

$$= \frac{-1 \pm \sqrt{1+600}}{2}$$

$$= -\frac{1}{2} - \frac{\sqrt{601}}{2} \quad \text{and} \quad -\frac{1}{2} + \frac{\sqrt{601}}{2}$$

$$= -12.76 \quad \text{and} \quad 11.76.$$

t is days, so we throw the solution of $t = -12.76$ out because it doesn't make sense. Our critical point is $t = 11.76$ days. To demonstrate that this point describes a local maximum, we use the second derivative test. the second derivative of $U(t)$ is

$$U''(t) = \frac{d}{dt}\left(\frac{30}{t+1} - \frac{t}{5}\right)$$

$$U''(t) = \frac{-30}{(t+1)^2} - \frac{1}{5}.$$

Plugging in $t = 11.76$, we get

$$U''(11.76) = \frac{-30}{(11.76+1)^2} - \frac{1}{5} = -.38.$$

So the second derivative test tells us that $t = 11.76$ is a local maximum. There are no additional critical points to consider, but t is bounded below by $t = 0$. The function at $t = 0$ is

$$U(0) = 30\ln(0+1) - \frac{0^2}{10}$$

$$= 30\ln(1) - 0 = 0.$$

The value of the function at $t = 11.76$ is

$$U(11.76) = 30\ln(11.76+1) - \frac{11.76^2}{10} = 62.6.$$

So $t = 11.76$ is also the global maximum. Therefore, the House Republicans maximize their utility of the shutdown if it lasts 11.76 days.

4. The sum of squared errors is the following function of β:

$$f(\beta) = \sum_{i=1}^{N} (y_i^2 - 2\beta x_i y_i + \beta^2 x_i^2).$$

In order to minimize this function, we first have to take the derivative with respect to β:

$$f'(\beta) = \frac{d}{d\beta}\left(\sum_{i=1}^{N} (y_i^2 - 2\beta x_i y_i + \beta^2 x_i^2) \right).$$

Notice that a summation $\sum$ is the same thing as a sum. Since derivatives break up over addition, we can rewrite the derivative of the sum as the sum of the derivatives of the addends:

$$f'(\beta) = \sum_{i=1}^{N} \frac{d}{d\beta}\left(y_i^2 - 2\beta x_i y_i + \beta^2 x_i^2 \right).$$

We treat y_i and x_i as constants, so the derivative is

$$f'(\beta) = \sum_{i=1}^{N} (-2x_i y_i + 2\beta x_i^2).$$

To find the critical point, we set the derivative equal to 0 and solve for β:

$$\sum_{i=1}^{N} (-2x_i y_i + 2\beta x_i^2) = 0.$$

We can rewrite this by taking the sum of each part:

$$\sum_{i=1}^{N} (-2x_i y_i) + \sum_{i=1}^{N} (2\beta x_i^2) = 0.$$

Factors without a subscript i can be brought outside the summations:

$$-2\sum_{i=1}^{N} x_i y_i + 2\beta \sum_{i=1}^{N} x_i^2 = 0.$$

And now we can simply solve for β:

$$\begin{aligned} 2\beta \sum_{i=1}^{N} x_i^2 &= 2\sum_{i=1}^{N} x_i y_i, \\ \beta \sum_{i=1}^{N} x_i^2 &= \sum_{i=1}^{N} x_i y_i, \\ \beta &= \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i^2}. \end{aligned}$$

Finally, to demonstrate that this critical point represents a local minimum, we find the second derivative of the sum of squares:

$$\begin{aligned} f''(\beta) &= \frac{d}{d\beta}\left(\sum_{i=1}^{N} (-2x_i y_i + 2\beta x_i^2) \right) \\ &= \sum_{i=1}^{N} \frac{d}{d\beta}\left(-2x_i y_i + 2\beta x_i^2 \right) \\ &= \sum_{i=1}^{N} (2x_i^2). \end{aligned}$$

The second derivative does not depend on β, and the since the x_i datapoints are squared, the sum $\sum_{i=1}^{N}(2x_i^2)$ must be positive. Therefore the critical point describes a local minimum.

5. (a) The trick here is to remember that we are taking the derivative with respect to μ. Notice that the first term in the log-likelihood function does not contain μ, so it is a constant and it drops out of the derivative. The derivative is

$$\ell'(\mu) = \frac{d}{d\mu}\Big(-.5\sum_{i=1}^{n}(x_i-\mu)^2\Big).$$

The constant factor -.5 comes in front of the derivative:

$$\ell'(\mu) = -.5\frac{d}{d\mu}\Big(\sum_{i=1}^{n}(x_i-\mu)^2\Big),$$

and the sum comes outside the derivative too since derivatives break up across addition:

$$\ell'(\mu) = -.5\sum_{i=1}^{n}\frac{d}{d\mu}(x_i-\mu)^2.$$

Here x_i is a constant and μ is the variable, so the derivative becomes

$$\ell'(\mu) = -.5\sum_{i=1}^{n}-2(x_i-\mu),$$

$$\ell'(\mu) = \sum_{i=1}^{n}(x_i-\mu).$$

To simplify this function, the summation can be applied to each term in the parentheses:

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\mu,$$

and since μ does not have an index, it is added n times. Therefore it can be rewritten as

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - n\mu.$$

(b) We've done almost all the work in part (a). We set the derivative equal to 0,

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - n\mu = 0,$$

and solve for μ:

$$\sum_{i=1}^{n}x_i = n\mu,$$

$$\mu = \frac{\sum_{i=1}^{n}x_i}{n}.$$

(c) The first derivative again is

$$\ell'(\mu) = \sum_{i=1}^{n}x_i - n\mu.$$

So the second derivative is just

$$\ell''(\mu) = -n.$$

That value is negative everywhere, so the value at the critical point we derived in part (b) is negative, and therefore the critical value refers to a local maximum. Since the domain of the normal distribution is unbounded, this value is also the global maximum.

(d) The critical point

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

is actually the mean of the sample of x values since it is the sum of the n values of x in the sample, divided by n. So we are estimating the mean of the normal distribution with the mean of the sample. This estimate makes perfect sense.