Math Camp Problem Set

Answer Key

Alex Houtz and Jacob Hukill*

August 2023

Day 1

1. **De Morgan's Laws** Let A and B both be events in Ω, the sample space.

   (a) Prove that $(A \cap B)^C = A^C \cup B^C$.

   (b) Prove that $(A \cup B)^C = A^C \cap B^C$.

Part (a) Solution:

Claim: $(A \cap B)^C = A^C \cup B^C$

Proof. Let $a \in (A \cap B)^C$. Then $a \in A \backslash B$, $B \backslash A$, or $A^C \cap B^C$. If $a \in A \backslash B$, then $a \in A \cap B^C$. So $a \in B^C$. If $a \in B \backslash A$, $a \in B \cap A^C$. So $a \in A^C$. If $a \in A^C \cap B^C$, $a \in B^C$ and $a \in A^C$. In all three cases, a is an element of either A^C, B^C, or both. Therefore, $a \in A^C \cup B^C$.

Now let $a \in A^C \cup B^C$. Then $a \in A^C$ or $a \in B^C$. If $a \in A^C$, $a \notin A$. So $a \notin A \cap B$. Therefore, $a \in (A \cap B)^C$. If $a \in B^C$, $a \notin B$. So $a \notin A \cap B$. Once again, $a \in (A \cap B)^C$. In both cases, $a \in (A \cap B)^C$.

We have shown that $(A \cap B)^C \subset A^C \cup B^C$ and that $A^C \cup B^C \subset (A \cap B)^C$. We therefore conclude that $(A \cap B)^C = A^C \cup B^C$. ■

*Math Camp 2023 Instructors | University of Notre Dame

Part (b) Solution:

Claim: $(A \cup B)^C = A^C \cap B^C$

Proof. Let $a \in (A \cup B)^C$. Then $a \notin A \cup B$. Therefore $a \notin A$ and $a \notin B$. So $a \in A^C$ and $a \in B^C$. We conclude that $a \in A^C \cap B^C$.

Now let $a \in A^C \cap B^C$. Then $a \in A^C$ and $a \in B^C$. So we know that $a \notin A$ and $a \notin B$. Therefore $a \notin A \cup B$ and we must conclude that $a \in (A \cup B)^C$.

We have shown that $(A \cup B)^C \subset A^C \cap B^C$ and that $A^C \cap B^C \subset (A \cup B)^C$. We therefore conclude that $(A \cup B)^C = A^C \cap B^C$. ■

2. **Hansen 1.3 Adapted** From a 52-card deck of playing cards, draw five cards to make a hand.

   (a) Let A be the event that the hand has exactly two Kings. Find $P(A)$.

   (b) Let B be the event the hand is a straight (not including straight flushes). Find $P(B)$.

   (c) Let C be the event that the hand is a flush (not including straight flushes). Find $P(C)$.

Part (a) Solution: Of the 13 values in a deck of cards, we can only choose one: the Kings. Of the 4 Kings in the deck, we only need to choose 2. The remaining 3 cards in our hand can be any of the remaining cards that are not Kings, minus any combinations that give us a full house or two pair. Mathematically:

$$
\begin{aligned}
P(\text{Two Kings}) &= \frac{1 \cdot \binom{4}{2}\left[\binom{48}{3} - \binom{12}{1}\binom{4}{3} - \binom{12}{1}\binom{4}{2}\binom{44}{1}\right]}{\binom{52}{5}} \\
&\approx 3.25\%
\end{aligned}
$$

Part (b) Solution: First, we list the possible number of straights we could draw (ignoring suit):

$$\begin{array}{ll}
A, 2, 3, 4, 5 & 2, 3, 4, 5, 6 \\
3, 4, 5, 6, 7 & 4, 5, 6, 7, 8 \\
5, 6, 7, 8, 9 & 6, 7, 8, 9, 10 \\
7, 8, 9, 10, J & 8, 9, 10, J, Q \\
9, 10, J, Q, K & 10, J, Q, K, A
\end{array}$$

So the starting card could be 1 of any 10 values. The value of the next 4 cards is then determined. Now, we just have to pick a suit for each card. This approach gives us the total number of straights.

But we want to exclude straight flushes. How many straight flushes are there? We can once again start with any of the 10 starting values. Only the first card, though, can be drawn from any suit. The next cards must then be that suit and the following value. Mathematically:

$$\begin{aligned}
P(\text{Straight}) &= \frac{\binom{10}{1}\left[\binom{4}{1}\right]^5 - \binom{10}{1}\binom{4}{1}\left[\binom{1}{1}\right]^4}{\binom{52}{5}} \\
&\approx 0.4\%
\end{aligned}$$

Part (c) Solution: For a flush, we don't care about the value of the cards. We only care about the suit. Therefore, we must first choose 1 of 4 suits. Then, of the 13 cards in that suit, we must choose 5. Lastly, we want to exclude drawing a straight flush. So:

$$P(\text{Flush}) = \frac{\binom{4}{1}\binom{13}{5} - \binom{10}{1}\binom{4}{1}\left[\binom{1}{1}\right]^4}{\binom{52}{5}}$$
$$\approx 0.2\%$$

This type of poker is called "five-card stud." So when playing five-card stud, know that the probability of drawing even a straight is less than half a percent.

3. **The Monte Hall Problem** You are on the game show "Let's Make a Deal with Monte Hall." There are three doors in front of you: doors A, B, and C. Your goal is to select the door with the prize behind it. Assume, without loss of generality, that you select door A. Monte then opens one of the other two doors, say door B, revealing that there is no prize behind it. He then gives you the option to switch your choice of doors. Should you stick with door A or switch your choice to door C? Assume that the *ex ante* probability of the prize being behind each door is 1/3.

Solution: First, we define three variables:

$$W \equiv \text{Door is the winning door}$$
$$S \equiv \text{Door selected by player}$$
$$H \equiv \text{Door opened by Monte}$$

Without loss of generality (WLOG), we pick door A. The probabilities of the prize being behind each door are:

$$
\begin{aligned}
P(W = A|S = A) &= \frac{1}{3} \\
P(W = B|S = A) &= \frac{1}{3} \\
P(W = C|S = A) &= \frac{1}{3}
\end{aligned}
$$

Now, WLOG, we assume that Monte opens door B. We have two probabilities to compute. Using Bayes' Rule:

$$
P(W = A|H = B, S = A) = \frac{P(H = B|W = A, S = A)P(W = A|S = A)}{P(H = B|S = A)} \tag{1}
$$

$$
P(W = C|H = B, S = A) = \frac{P(H = B|W = C, S = A)P(W = C|S = A)}{P(H = B|S = A)} \tag{2}
$$

Let's look at the denominator using the law of total probability:

$$
\begin{aligned}
P(H = B|S = A) = P(H = B|W = A, S = A) \cdot P(W = A|S = A) \\
+ P(H = B|W = B, S = A) \cdot P(W = B|S = A) \\
+ P(H = B|W = C, S = A) \cdot P(W = C|S = A)
\end{aligned}
$$

Above, we listed all probabilities of the form $P(W = w|S = A)$. What about the remaining probabilities of the form $P(H = B|W = w, S = A)$? We know each of these as well by thinking about the problem:

$$
\begin{aligned}
P(H = B|W = A, S = A) &= \frac{1}{2} \\
P(H = B|W = B, S = A) &= 0 \\
P(H = B|W = C, S = A) &= 1
\end{aligned}
$$

Now we can input each of these probabilities into equation (1):

$$
\begin{aligned}
P(W = A|H = B, S = A) &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} \\
&= \frac{\frac{1}{2}}{\frac{1}{2} + 1}
\end{aligned}
$$

$$= \frac{1}{3} \tag{1}$$

and into equation (2):

$$\begin{aligned} P(W = A|H = B, S = A) &= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} \\ &= \frac{1}{\frac{1}{2} + 1} \\ &= \frac{2}{3} \end{aligned} \tag{2}$$

Therefore, the probability that the prize behind the door you first choose is $\frac{1}{3}$, while the probability that the prize is behind the door you could switch to is $\frac{2}{3}$. If you want a greater likelihood of winning, switch your door choice.

4. **The St. Petersburg Paradox** Suppose a wealthy billionaire runs a game. Each participant gets to flip a fair coin until the coin comes up as heads for the first time. At that point, the participant wins $\$2^n$, where n denotes the number of flips. The billionaire charges \$1 billion to play.

   (a) Write down the pmf and verify that it is valid.

   (b) What is the probability that you win more than \$4?

   (c) Suppose you are an economic agent that makes decisions based only on your expected payoff. Should you pay the entry fee?

Part (a) Solution: Let N denote the number of flips until H appears. The probability $N = 1$ is $\frac{1}{2}$, as the $P(H) = \frac{1}{2}$ on the first flip. The probability $N = 2$ is $\frac{1}{4}$, as a T must be flipped on the first toss and then a H must be flipped on the second toss. The probability $N = 3$ is $\frac{1}{8}$, because a T must be flipped on both the first and second flip to reach the third flip, where a H must be flipped. If we continue on, we see that:

$$f_N(n) = \left(\frac{1}{2}\right)^n$$

To prove this is a valid pmf, we must show that it sums to 1 over all values of n. So:

$$\begin{aligned}
\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n &= \frac{\frac{1}{2}}{1 - \frac{1}{2}} \\
&= \frac{\frac{1}{2}}{\frac{1}{2}} \\
&= 1
\end{aligned}$$

We used the sum of an infinite geometric series to cross the first equal sign. We must also show that for each value of n, $f(n) > 0$. But because n must be a positive integer, this requirement is also satisfied.

Part (b) Solution: The payoff is defined as \$$2^n$. So:

$$\begin{aligned}
4 &= 2^n \\
log_2(4) &= n \\
2 &= n
\end{aligned}$$

We are looking at the probability that $N > 2$. Then:

$$\begin{aligned}
P(N > 2) &= 1 - P(N \leq 2) \\
&= 1 - \sum_{n=1}^{2} \left(\frac{1}{2}\right)^n \\
&= 1 - \frac{1}{2} + \frac{1}{4} \\
&= 1 - \frac{3}{4} \\
&= \frac{1}{4}
\end{aligned}$$

Part (c) Solution: Let $W = 2^n$ denote your winnings. Then:

$$\begin{aligned}
\mathbb{E}[W] &= \sum_{n=1}^{\infty} 2^n \left(\frac{1}{2}\right)^n \\
&= \sum_{n=1}^{\infty} \frac{2^n \cdot 1^n}{2^n} \\
&= \sum_{n=1}^{\infty} 1 \\
&= \infty
\end{aligned}$$

Because the expected payoff is infinite, an agent that only considers the expected payoff should pay the $1 billion to play.

5. **Hansen 2.6 Edited** Compute $\mathbb{E}[X]$ and $Var(X)$ for the following distributions:

   (a) $f(x) = ax^{-a-1}$, for $1 < x < \infty$ and $a > 0$.

   (b) $f(x) = \frac{1}{n}$, for $x = 1, 2, ..., n$. Hint: show by induction that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ and that $\sum_{i=1}^{n} i^2 = \frac{n(2n+1)(n+1)}{6}$.

   (c) $f(x) = \frac{3}{2}(x-1)^2$, for $0 < x < 2$.

Part (a) Solution:

This is a continuous distribution, so we use the integral definition of the expected value:

$$\begin{aligned}
\mathbb{E}[X] &= \int_1^{\infty} xax^{-a-1} dx \\
&= a \int_1^{\infty} x^{-a} dx \\
&= a \left[\frac{1}{1-a} x^{-a+1}\right]_1^{\infty} \\
&= \frac{a}{a-1}
\end{aligned}$$

To calculate the variance, we want to find $\mathbb{E}[X^2]$ first:

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_1^\infty x^2 a x^{-a-1} dx \\
&= a \int_1^\infty x^{-a+1} dx \\
&= a \left[\frac{1}{2-a} x^{-a+2} \right]_1^\infty \\
&= \frac{a}{a-2}
\end{aligned}$$

So the variance is:

$$\begin{aligned}
Var(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \frac{a}{a-2} - \frac{a^2}{(a-1)^2}
\end{aligned}$$

<u>**Part (b) Solution:**</u> This distribution is discrete, so we use the summation definition of the expectation:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=1}^n i \cdot \frac{1}{n} \\
&= \frac{1}{n} \sum_{i=1}^n i \qquad (*) \\
&= \frac{1}{n} \frac{n(n+1)}{2} \qquad (**) \\
&= \frac{n+1}{2}
\end{aligned}$$

To get from equation $(*)$ to equation $(**)$, we use a proof by induction (below).

<u>Claim:</u> $\sum_{i=1}^n i = \frac{n(n+1)}{2}$

Proof (by Induction). **Base Case:** Let $n = 1$. Then:

$$\sum_{i=1}^{n} i = \sum_{i=1}^{1} i = 1$$

Similarly:

$$\frac{n(n+1)}{2} = \frac{1(1+1)}{2} = 1$$

So the claim holds.

Inductive Step: Assume that for all values less than n, the claim is true. Then:

$$\begin{aligned}
\sum_{i=1}^{n-1} i &= \frac{(n-1)(n-1+1)}{2} \\
1 + 2 + ... + (n-1) + n &= \frac{(n-1)(n-1+1)}{2} + n \\
\sum_{i=1}^{n} i &= \frac{(n-1)n}{2} + \frac{2n}{2} \\
&= \frac{n^2 - n + 2n}{2}
\end{aligned}$$

$$\begin{aligned}
&= \frac{n^2 + n}{2} \\
&= \frac{n(n+1)}{2}
\end{aligned}$$

This then proves the claim. ■

Now we can calculate the second moment:

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{i=1}^{n} i^2 \cdot \frac{1}{n} \\
&= \frac{1}{n} \sum_{i=1}^{n} i^2 && (\#) \\
&= \frac{1}{n} \frac{n(n+1)(2n+1)}{6} && (\#\#)
\end{aligned}$$

$$= \frac{(n+1)(2n+1)}{6}$$

To get from (#) to (##), we use another proof by induction (below).

Claim: $\sum_{i=1}^{n} i^2 = \frac{n(2n+1)(n+1)}{6}$

Proof (by Induction). **Base Case:** Let $n = 1$. Then:

$$\sum_{i=1}^{n} i^2 = \sum_{i=1}^{1} i^2 = 1$$

Likewise:

$$\frac{n(2n+1)(n+1)}{6} = \frac{(3)(2)}{6} = 1$$

So the claim holds.

Inductive Step: Now assume that the claim holds for all values less than n. Then:

$$\begin{aligned}
\sum_{i=1}^{n-1} i^2 &= \frac{(n-1)(2(n-1)+1)(n-1+1)}{6} \\
1 + 4 + ... + (n-1)^2 + n^2 &= \frac{(n-1)(2(n-1)+1)(n-1+1)}{6} + n^2 \\
\sum_{i=1}^{n} i^2 &= \frac{2n^3 - 2n^2 - n^2 + n + 6n^2}{6} \\
&= \frac{(2n^3 + 3n^2 + n}{6} \\
&= \frac{n(2n+1)(n+1)}{6}
\end{aligned}$$

But this proves the claim. ■

Lastly, we can calculate the variance:

$$\begin{aligned} Var(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \frac{(n+1)(2n+1)}{6} - \frac{n^2(n+1)^2}{4} \end{aligned}$$

<u>**Part (c) Solution:**</u> This variable is continuous again, so we use the integral definition of the expectation:

$$\begin{aligned} \mathbb{E}[X] &= \int_0^2 x\frac{3}{2}(x-1)^2 dx \\ &= \frac{3}{2}\int_0^2 x(x^2-2x+1)dx \\ &= \frac{3}{2}\int_0^2 x^3 - 2x^2 + xdx \\ &= \frac{3}{2}\left[\frac{1}{4}x^4 - \frac{2}{3}x^3 + \frac{1}{2}x^2\right]_0^2 \end{aligned}$$

$$\begin{aligned} &= \frac{3}{2}\left(4 - \frac{16}{3} + 2\right) \\ &= 1 \end{aligned}$$

Similarly, we can find the second moment:

$$\begin{aligned} \mathbb{E}[X^2] &= \int_0^2 x^2\frac{3}{2}(x-1)^2 dx \\ &= \frac{3}{2}\int_0^2 x^2(x^2-2x+1)dx \\ &= \frac{3}{2}\int_0^2 x^4 - 2x^3 + x^2dx \\ &= \frac{3}{2}\left[\frac{1}{5}x^5 - \frac{1}{2}x^4 + \frac{1}{3}x^3\right]_0^2 \\ &= \frac{3}{2}\left(\frac{32}{5} - 8 + \frac{8}{3}\right) \end{aligned}$$

$$= \frac{8}{5}$$

6. **Hansen 2.7** Let X have density:

$$f_X(x) = \frac{1}{2^{r/2}\Gamma(r/2)} x^{r/2-1} e^{-x/2}$$

for $x > 0$. Let $Y = 1/X$. Derive the density of Y for $y > 0$.

<u>**Solution:**</u> Let's look at the change-of-variables equation first:

$$f_Y(y) = \left| \frac{dh^{-1}(y)}{dy} \right| f_X(h^{-1}(y))$$

Let's find h^{-1}:

$$\begin{aligned} Y &= h(x) \\ &= \frac{1}{X} \end{aligned} \qquad\qquad \begin{aligned} X &= h^{-1}(Y) \\ &= \frac{1}{Y} \end{aligned}$$

Next, plug h^{-1} into f_X:

$$f_X\left(\frac{1}{y}\right) = \frac{1}{2^{r/2}\Gamma(r/2)} \left(\frac{1}{y}\right)^{r/2-1} e^{-1/(2y)}$$

Now find the Jacobian:

$$\begin{aligned} \left| \frac{dh^{-1}(y)}{dy} \right| &= \left| -\frac{1}{y^2} \right| \\ &= \frac{1}{y^2} \end{aligned}$$

Altogether:

$$
\begin{aligned}
f_Y(y) &= \frac{1}{y^2}\frac{1}{2^{r/2}\Gamma(r/2)}\left(\frac{1}{y}\right)^{r/2-1} e^{-1/(2y)} \\
&= \frac{1}{2^{r/2}\Gamma(r/2)}\left(\frac{1}{y}\right)^{r/2+1} e^{-1/(2y)} \\
&= \frac{1}{2^{r/2}\Gamma(r/2)} y^{-r/2-1} e^{-1/(2y)}
\end{aligned}
$$

7. **Hansen 4.3** Let:

$$
f(x,y) = \frac{2}{(1+x+y)^3}
$$

for $x \geq 0$ and $0 \leq y$.

(a) Verify that $f(x,y)$ is a valid density.

(b) Find $f_x(x)$

(c) Find $\mathbb{E}[y]$, $Var(y)$, $\mathbb{E}[xy]$, and $\rho_{x,y}$

(d) Find $f(y|x)$

(e) Find $\mathbb{E}[y|x]$

Part (a) Solution: We show that $f(x,y)$ integrates to one:

$$
\begin{aligned}
\int_0^\infty \int_0^\infty 2(1+x+y)^{-3} dx dy &= \int_0^\infty -(1+x+y)^{-2}\Big]_0^\infty dy \\
&= \int_0^\infty (1+y)^{-2} dy \\
&= -(1+y)^{-1}\Big]_0^\infty \\
&= 1
\end{aligned}
$$

Part (b) Solution: We integrate y out of the joint pdf:

$$
\begin{aligned}
f(x) &= \int_0^\infty 2(1+x+y)^{-3} dy \\
&= -(1+x+y)^{-2}\Big]_0^\infty \\
&= (1+x)^{-2}
\end{aligned}
$$

Part (c) Solution: Using the definition of the expectation, and recognizing that the marginal densities are symmetric:

$$
\begin{aligned}
\mathbb{E}[y] &= \int_0^\infty y f(y) dy \\
&= \int_0^\infty y(1+y)^{-2} dy
\end{aligned}
$$

Use u-substitution, defining $u = 1 + y$ and $du = dy$:

$$
\begin{aligned}
\mathbb{E}[y] &= \int_1^\infty \frac{u-1}{u^2} du \\
&= \int_1^\infty \frac{1}{u} - \frac{1}{u^2} du \\
&= \ln|u|\Big]_1^\infty + \frac{1}{u}\Big]_1^\infty \\
&\to \infty
\end{aligned}
$$

The expectation is divergent. Therefore, the variance is also divergent:

$$
\begin{aligned}
Var(y) &= \mathbb{E}[y^2] - \mathbb{E}[y]^2 \\
&= \int_0^\infty y^2(1+y)^{-2} dy - \left(\lim_{y\to\infty}\left[\ln|1+y| + (1+y)^{-1}\right]\right)^2 \\
&= \int_1^\infty \frac{(u-1)^2}{u^2} du - \left(\lim_{y\to\infty}\left[\ln|1+y| + (1+y)^{-1}\right]\right)^2 \\
&= \int_1^\infty 1 - \frac{2}{u} + \frac{1}{u^2} du - \left(\lim_{y\to\infty}\left[\ln|1+y| + (1+y)^{-1}\right]\right)^2
\end{aligned}
$$

$$= \left[u - 2\ln|u| - \frac{1}{u}\right]_1^\infty - \left(\lim_{y\to\infty} \left[\ln|1+y| + (1+y)^{-1}\right]\right)^2$$
$$\to \infty$$

where $u = 1 + y$ and $du = dy$. What about the joint expectation?

$$\begin{aligned}
\mathbb{E}[xy] &= \int_0^\infty \int_0^\infty 2xy(1+x+y)^{-3}dxdy \\
&= \int_0^\infty 2y \int_0^\infty x(1+x+y)^{-3}dxdy
\end{aligned}$$

Let $u = 1 + x + y$ and $du = dx$:

$$\begin{aligned}
\mathbb{E}[xy] &= \int_0^\infty 2y \int_{1+y}^\infty \frac{u-y-1}{u^3}dudy \\
&= \int_0^\infty 2y \left[-\frac{1}{u^2} + \frac{1+y}{2u^2}\right]_{1+y}^\infty dy \\
&= \int_0^\infty 2y \left[\frac{1}{1+y} - \frac{1+y}{2(1+y)^2}\right] dy \\
&= \int_0^\infty y(1+y)^{-1}dy
\end{aligned}$$

Now let $w = 1 + y$ and $dw = dy$:

$$\begin{aligned}
\mathbb{E}[xy] &= \int_1^\infty \frac{w-1}{w}dw \\
&= [w - \ln|w|]_1^\infty \\
&\to \infty
\end{aligned}$$

Since the joint expectation diverges, so does the correlation.

Part (d) Solution: Using the conditional probability formula:

$$\begin{aligned}
f(y|x) &= \frac{f(x,y)}{f(x)} \\
&= \frac{2(1+x+y)^{-3}}{(1+x)^{-2}}
\end{aligned}$$

$$= \frac{2(1+x)^2}{(1+x+y)^3}$$

Part (e) Solution: Using the definition of the expectation again:

$$\begin{aligned}
\mathbb{E}[y|x] &= \int_0^\infty y \frac{2(1+x)^2}{(1+x+y)^3} dy \\
&= 2(1+x)^2 \int_0^\infty \frac{y}{(1+x+y)^3} dy
\end{aligned}$$

Let $u = 1 + x + y$ and $du = dx$:

$$\begin{aligned}
\mathbb{E}[y|x] &= 2(1+x)^2 \int_{1+x}^\infty \frac{u-x-1}{u^3} du \\
&= 2(1+x)^2 \int_{1+x}^{infty} \frac{1}{u^2} - \frac{1+x}{u^3} du \\
&= 2(1+x)^2 \left[-\frac{1}{u} + \frac{1+x}{2u^2} \right]_{1+x}^\infty \\
&= 2(1+x)^2 \left[\frac{1}{1+x} - \frac{1+x}{2(1+x)^2} \right] \\
&= 2(1+x) - (1+x) \\
&= 1 + x
\end{aligned}$$

8. $\boldsymbol{\varepsilon - \delta}$ **Continuity** Let $f(x) = \frac{1}{x}$. Prove that for $x \neq 0$, $f(x)$ is continuous (Hint: you may have to assume that $\delta < \frac{1}{2}|x_0|$ to solve for the appropriate δ).

Solution:

Claim: If $f(x) = \frac{1}{x}$, then $f(x)$ is continuous for $x \neq 0$.

First, suppose the claim is true. Then:

$$\begin{aligned}
|f(x) - f(x_0)| &< \varepsilon \\
\left| \frac{1}{x} - \frac{1}{x_0} \right| &< \varepsilon
\end{aligned}$$

$$\begin{aligned}
\left| \frac{x - x_0}{xx_0} \right| &< \varepsilon \\
\frac{|x - x_0|}{|x||x_0|} &< \varepsilon \\
\frac{\delta}{|x||x_0|} &< \varepsilon
\end{aligned}$$

Now, we have $|x|$ in the problem. We need to substitute it out. Let's assume that:

$$\begin{aligned}
|x - x_0| &< \frac{1}{2}|x_0| \\
\implies |x| &< \frac{3}{2}|x_0|
\end{aligned}$$

Then:

$$\begin{aligned}
\frac{\delta}{\frac{3}{2}|x_0|^2} &< \varepsilon \\
\frac{2\delta}{3|x_0|^2} &< \varepsilon \\
\delta &< \frac{3}{2}\varepsilon|x_0|^2
\end{aligned}$$

Now we can commence with the proof:

Proof. Let $\varepsilon > 0$ and $|x - x_0| < \delta$. Pick δ such that $\delta < \frac{1}{2}|x_0|$ and $\delta < \frac{3}{2}\varepsilon|x_0|^2$. Then:

$$\begin{aligned}
|x - x_0| &< \delta \\
|x - x_0| &< \frac{3}{2}\varepsilon|x_0|^2
\end{aligned}$$

$$\frac{2}{3}\frac{|x-x_0|}{|x_0|^2} < \varepsilon$$
$$\frac{|x-x_0|}{|x||x_0|} < \varepsilon$$
$$\left|\frac{x-x_0}{xx_0}\right| < \varepsilon$$
$$\left|\frac{1}{x}-\frac{1}{x_0}\right| < \varepsilon$$

But this is just:

$$|f(x)-f(x_0)| < \varepsilon$$

This proves continuity. ■

9. Is the intersection of an infinite collection of open sets open? Either prove the claim, or find a counterexample.

   <u>**Solution**</u>: Let the family of open intervales $\{A_n\}$ be given by

   $$A_n = (-1/n, 1/n)$$

   This forms a sequence of open intervals:

   $$\left\{(-1,1), (\frac{-1}{2}, \frac{1}{2}, ...\right\}.$$

   Then as $n \to \infty$, it is clear that $A_n \to \{0\}$. Therefore:

   $$\bigcap_{n=1}^{\infty} A_n = \{0\}$$

   which is not open (it is actually closed).

10. **Compactness of a Budget Set** A type of compact set that you will encounter often is called a budget set. Consider a vector of positive prices $p \in \mathbb{R}^N, p \gg 0$ and a wealth level $w \geq 0$. The budget set contains all affordable bundles of goods, and is given by

    $$B(p,w) = \{x \in \mathbb{R}^N | x \geq 0 \text{ and } p \cdot x \leq w\}$$

In two dimensions, this looks like a triangle with vertices at $0, w/p_1, w/p_2$. Show that this set is compact.

Solution: Suppose that $\varepsilon = max_i \ w/\_i + 1$. Note that $B(p,w) \subseteq B_\varepsilon(0)$. This proves that $B(p,w)$ is bounded. Now consider any $x \in \mathbb{R}^N \diagdown B(p,w)$. Let

$$\varepsilon = min \left\{ \{-x_i | x_i < 0\} \cup \sqrt{\frac{w - p \cdot x}{p \cdot p}} \right\} > 0,$$

and note that $B_\varepsilon(x) \subseteq \mathbb{R}^N \diagdown B(p,w)$. This shows that the complement of $B(p,w)$ is open, and hence $B(p,w)$ is closed. Therefore $B(p,w)$ is both closed and open, and thus compact.

11. **Taylor Approximation of a Quadratic** Let $f(x) = ax^2 + bx + c$. Write the second order Taylor approximation around 0 for the function. What happens?

    Solution: Let's start with the derivatives: $f'(x) = 2ax + b$ and $f''(x) = 2a$. Then the Taylor series around 0 is given by

$$\begin{aligned} f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 &= c + bx + \frac{2a}{2!}x^2 \\ &= ax^2 + bx + c \end{aligned}$$

    Does this surprise you? If we approximate the quadratic function with another quadratic function, then we will reconstruct the given function completely.

Day 2

12. **Positive Semi-Definite Matrices** Suppose we have:

$$C = (x'x)^{-1}(x'\Omega x)(x'x)^{-1} - (x'\Omega^{-1}x)^{-1}$$

    Suppose that Ω is positive definite such that a Cholesky decomposition exists. Prove, using the following steps, that C is positive semi-definite. Note that $A - B$ is positive semi-definite iff $B^{-1} - A^{-1}$ is also positive semi-definite.

Solution:

Proof. The problem tells us that $A - B$ is positive semi-definite iff $B^{-1} - A^{-1}$ is positive semi-definite. Therefore, let's first invert A and B:

$$\begin{aligned} A^{-1} &= \left[(x'x)^{-1}(x'\Omega x)(x'x)^{-1}\right]^{-1} \\ &= (x'x)(x'\Omega x)^{-1}(x'x) \\ B^{-1} &= \left[(x'\Omega^{-1}x)^{-1}\right]^{-1} \end{aligned}$$

$$= (x'\Omega^{-1}x)$$

Then $B^{-1} - A^{-1}$ is:

$$B^{-1} - A^{-1} = (x'\Omega^{-1}x) - (x'x)(x'\Omega x)^{-1}(x'x)$$

Factor out an x' to the front and an x to the back:

$$B^{-1} - A^{-1} = x'\left[\Omega^{-1} - x(x'\Omega x)^{-1}x'\right]x$$

We have reached the expression given in part (a). Now Cholesky decompose Ω:

$$\begin{aligned} B^{-1} - A^{-1} &= x'\left[\Omega^{-1/2}\Omega^{-1/2'} - x\left(x'\Omega^{1/2}\Omega^{1/2'}x\right)^{-1}x'\right]x \\ &= x'\Omega^{-1/2}\left[I - x\Omega^{1/2}\left(x'\Omega^{1/2}\Omega^{1/2'}x\right)^{-1}x'\Omega^{1/2'}\right]\Omega^{-1/2'}x \end{aligned}$$

Define $z = \Omega^{1/2}x$:

$$B^{-1} - A^{-1} = x'\Omega^{-1/2}\left[I - z(z'z)^{-1}z'\right]\Omega^{-1/2'}x$$

We see that $z(z'z)^{-1}z$ is the projection matrix and $I - z(z'z)^{-1}z = M_z$ is the annihilator matrix. The annihilator matrix is idempotent, so:

$$B^{-1} - A^{-1} = WM_zM_zW'$$

where $W = x'\Omega^{-1/2}$. M_z is also symmetric:

$$\begin{aligned} B^{-1} - A^{-1} &= W M_z M_z' W' \\ &= \|W M_z\|^2 \\ &\geq 0 \end{aligned}$$

Therefore, $B^{-1} - A^{-1}$ is PSD, which implies that $A - B$ is PSD, proving the claim. ■

13. Consider the problem[1] of maximizing $f(x, y) = x$ subject to the constraint $x^3 + y^2 = 0$.

    - Try using the Lagrangian method to find the solution. What happens? (Note: you do not have to verify that the point you find is indeed a maximum)

    Solution:
    We set up the Lagrangian:

    $$L(x, y, \lambda) = x - \lambda(x^3 + y^2 - 0)$$

    We then take our first order conditions to solve the system. Notice that we have an equality constraint, so we don't need the complementary slackness conditions:

    $$\begin{aligned} \frac{\partial L}{\partial x} &= 1 - 3\lambda x^2 = 0 \\ \frac{\partial L}{\partial y} &= -2\lambda y = 0 \\ \frac{\partial L}{\partial \lambda} &= -(x^3 + y^2) = 0 \end{aligned}$$

    Notice that this system has no solution! From the first equation, neither x nor lambda can be zero. From the second equation, this ensures that y must be zero, since lambda cannot be zero. However, this creates a contradiction in the third equation. If y is zero, then x must also be zero for (3) to be true. But we have already said that x cannot be zero! Thus, this system has no solution.

[1] Note: Most problems are taken from Simon and Blume 2004

So where did we go wrong? Remember our theorem for us to use the Lagrangian. We must ensure that our constraint qualification is met, which we did not do. Let's do that now. The critical points of h are:

$$\frac{\partial h}{\partial x} = 3x^2 = 0$$
$$\frac{\partial h}{\partial y} = 2y = 0$$

The only solution here is (0,0), which does lie within the constraint set. Therefore, the only solution candidate of the maximization problem is the critical point of h, (0,0), which is the solution to the problem[2].

14. Consider the following maximization problem:

$$Maximize \;\; f(x,y) = x^3 + y^3$$
$$subject\ to \;\; x - y = 0$$

- Write down the Lagrangian, and give the FONCs for a maximum
- Solve the problem using the FONCs. What happened? Why might be the problem?

Solution:

We first check the constraint qualification, and see that it is satisfied, as the only critical point of g is (1,-1). Now, we write down the Lagrangian and the FONCs

$$L = x^3 + y^3 - \lambda(x - y)$$
$$\frac{\partial L}{\partial x} = 3x^2 - \lambda = 0$$
$$\frac{\partial L}{\partial y} = 3y^2 + \lambda = 0$$
$$\frac{\partial L}{\partial \lambda} x - y = 0$$

Now we can use the FONCs to try and solve the problem. From the first two

[2]You may notice that the constraint is not bounded, but note the following. The term y^2 can only be positive. So for the constraint to be satisfied, it has to be the case that x is either negative, or zero. It is clearly not maximal for x to be negative however.

equations, we can see that the only solution is x=y=λ=0. While this does satisfy the necessary conditions, it's clear that this cannot be a local max. In the neighborhood of x=y=0, we have some $x+\epsilon = y+\epsilon = \epsilon$, which satisfies the constraint, and makes the objective function larger than our candidate point.

This shows that our candidate is not a local constrained max, and certainly doesn't maximize the problem. What's the issue? Well, Lagrange's Theorem offers only the necessary conditions for local constrained optima, not sufficient conditions. It may be the case that there does not exist a global constrained max in the first place. In fact, we can see that our set is not compact, so we cannot use Weierstrass's theorem to guarantee that a max exists. So when you approach optimization problems, you may want to check the conditions to ensure that an optimum exists.

15. Suppose we want to maximize $f(x_1, x_2) = x_1 x_2$ subject to the constraint $x_1 + 4x_2 = 16$.

    - Write down the Lagrangian for this problem and solve it.

Solution:

We begin by checking the constraint qualification. The gradient of h is (1,4), meaning the constraint qualification is satisfied as h has no critical points. We can now form the Lagrangian:

$$L(x, y, \lambda) = x_1 x_2 - \lambda(x_1 + 4x_2 - 16)$$

We take the first order conditions by setting the partials equal to zero:

$$\frac{\partial L}{\partial x_1} = x_2 - \lambda = 0$$
$$\frac{\partial L}{\partial x_2} = x_1 - 4\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = -(x_1 + 4x_2 - 16) = 0$$

Note that we can only take the first order conditions like this, because we have an equality constraint. To solve this system, we find lambda in the first equation, and plug it into the second. We can then find that $x_1 = 4x_2$. We now plug this into the last equation, to get that $x_2 = 2$. Then the solution to this system is:

$$x_1 = 8, \quad x_2 = 2, \quad \lambda = 2$$

Our theorem states that the only candidate for a solution to this problem must be

$$x_1 = 8, \quad x_2 = 2$$

Note that it may not be immediate from the theorem that the point is a global or even local constrained maximum. However, it is the only point that satisfies the necessary conditions for a local constrained maximum. But the theorem we used gave necessary conditions, not sufficient ones. However, if we can prove that there does exist a global constrained maximum, then this global maximum will automatically be a local constrained maximum, which must also satisfy the necessary conditions. Given, that we only have one point in which they are satisfied, it would have to be the candidate above. We can in fact show that there does exist a global constrained maximum using Weierstrass's theorem. Our constraint set is closed and bounded, so our function must have a maximum on it. Thus, we can conclude that our candidate is the maximum.

16. Consider the following maximization problem:

$$\begin{aligned} &\textit{maximize} \quad f(x_1, x_2) = x_1^2 x_2 \\ &\textit{subject to} \quad 2x_1^2 + x_2^2 = 3 \end{aligned}$$

- Write down the Lagrangian and the FONCs for this problem.
- Without using the second order conditions, how would you find the constrained max? Do so.
- Now use the second order test to find the max.

Solution:

We start by checking the constraint qualification. The gradient of h is $(4x_1, 2x_2)$, so the only critical point is at (0,0). This point, however, is not in the constraint set, so the the constraint qualification is satisfied. We can now set up the Lagrangian and take first order conditions:

$$L(x_1, x_2, \lambda) = x_1^2 x_2 - \lambda(2x_1^2 + x_2^2 - 3)$$

$$\begin{aligned}
\frac{\partial L}{\partial x_1} &= 2x_1x_2 - 4\lambda x_1 = 0 \\
\frac{\partial L}{\partial x_2} &= x_1^2 - 2\lambda x_2 = 0 \\
\frac{\partial L}{\partial \lambda} &= -(2x_1^2 + x_2^2 - 3) = 0
\end{aligned}$$

We have a problem of three equations and three unknowns, but it is now a non-linear system, unlike those we've worked with so far. We'll need to do a little more work to get to the answer. Let's look at rewrite equation one:

$$\begin{aligned}
2x_1x_2 - 4\lambda x_1 &= 0 \\
2x_1(x_2 - 2\lambda) &= 0
\end{aligned}$$

From this, it is clear that either $x_1 = 0$, or $x_2 = 2\lambda$ for the condition to be satisfied. To solve this, we approach each case separately.

If $x_1 = 0$, then the third equation implies that either $x_2 = \sqrt{3}$, or $x_2 = -\sqrt{3}$. Furthermore, from the second equation it must be that $\lambda = 0$. Then we have two candidates for the solution:

$$(0, \sqrt{3}, 0), \quad (0, -\sqrt{3}, 0)$$

If $x_2 = 2\lambda$, then we can plug this value for λ into the second equation to get that $x_1^2 = x_2^2$. Now using the third equation, we find that either $x_1 = 1$ or $x_1 = -1$. Using our derived relations, we get four more candidates for the solution:

$$(1, 1, \frac{1}{2}), \quad (1, -1, \frac{-1}{2}), \quad (-1, 1, \frac{1}{2}), \quad (-1, -1, \frac{-1}{2})$$

We know that because we are maximizng a continuous function over a compact constraint, that the problem does have a constrained max. Furthermore, we know from our necessary conditions that our constrained max must be one of these candidates. We can evaluate each of these points by plugging them into the original objective function to find which gives the highest value.

$$f(0,\sqrt{3}) = f(0,-\sqrt{3}) = 0$$
$$f(1,1) = f(-1,1) = 1$$
$$f(1,-1) = f(-1,-1) = -1$$

So we can see that the max occurs at points (1,1) and (-1,1) (and that the min occurs at points (1,-1),(-1,-1).)

Let's try and apply the second order conditions by creating the bordered Hessian.

$$\begin{bmatrix} 0 & 4x_1 & 2x_2 \\ 4x_1 & 2x_2 - 4\lambda & 2x_1 \\ 2x_2 & 2x_1 & -2\lambda \end{bmatrix}$$

Now, for a local constrained max, we check the last n-(e+k) principal minors, with the largest matrix having the same sign as $(-1)^n$. This means we only need to check the principal minor of the whole matrix, and ensure that it is positive at the candidate. The determinant of the matrix is detH=-16 at candidates $(1,-1,\frac{-1}{2})$ and $(-1,-1,\frac{-1}{2})$. Thus, both of these points are local min. The determinant of the matrix is detH=48 at candidates $(1,1,\frac{1}{2})$ and $(-1,1,\frac{1}{2})$, so these points are both local max, and are indeed both the constrained max for this problem.

You might wonder why we would use the second order test, when in this case it was easier to just plug our values into the function and evaluate. I've previously talked about how often we aren't necessarily interested in exact values, but rather want to "characterize" the solution. Suppose that we don't have numbers in our problem, but instead have parameters with unknown values. We can't just plug the candidates into the function and evaluate. However, we can see for which values of these parameters will our second order condition hold for a max.

Furthermore, notice that we could not characterize the nature of the points $(0, \pm\sqrt{3}, 0)$ above. However, if we look at the determinant of H for these candidates, we can see that detH $= -24\sqrt{3}$ for point $4(0, \sqrt{3})$, and detH $= 24\sqrt{3}$ for point $(0, -\sqrt{3})$. This means the first point is a local min, and the second point is a local max. While we are often interested in local optima, this highlights an important point. **Note:** the points calculated using this second order test need not be **global** extrema, but may just be local. We can't prove that we have a global maximum x of a nonlinear function using derivatives at x alone; if we are interested in this, we need to prove global concavity of F within the feasible set, or use a different method to argue for it.

17. Consider the the problem of maximizing $f(x, y) = xy$ subject to the constraint g(x,y)$=x^2 + y^2 \leq 1$.

    - Form the Lagrangian and write out the FONCs.
    - How many candidates do you need to check?
    - Find the point that maximizes the problem

    Solution:
    The first step is to check the constraint qualification. We need to ensure that it holds whenever g is binding. Notice that the only critical point of g occurs at (0,0), away from the boundary, meaning that the constraint qualification holds.

    The Lagrangian is

    $$L(x, y, \lambda) = xy - \lambda(x^2 + y^2 - 1)$$

    and the first order conditions are

    $$\begin{aligned}
    &\frac{\partial L}{\partial x} = y - 2\lambda x = 0 \\
    &\frac{\partial L}{\partial y} = x - 2\lambda y = 0 \\
    &\lambda \geq 0 \\
    &x^2 + y^2 \leq 1 \\
    &\lambda(x^2 + y^2 - 1) = 0
    \end{aligned}$$

Remember that since we are dealing with an inequality constraint now, that we need our complementary slackness conditions instead of just writing down the constraint. Now, we can check cases. We start when lambda is zero.

If $\lambda = 0$, then x=y=0 from the first two equations. Because this satisfies all of the first order conditions, we consider it as a candidate for the solution.

Now let $\lambda > 0$. Then from the first two equations, $y/2x = x/2y$ implies $y = x$. Because lambda is not zero, then we know that the constraint must bind with equality. Then $2x^2 = 1$ implies that $x = y = \pm\frac{1}{\sqrt{2}}$. We can combine these values for x and y to find lambda from the first two equations, and get the following candidates:

$$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{2}), \quad (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{-1}{2}), \quad (\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{-1}{2}), \quad (\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{1}{2})$$

Notice that in the middle two cases, this generates negative values for the multiplier lambda. This gives us a contradiction, as we require that lambda be positive. Given this is the case, we can disregard these candidates, and are only left with three candidates that we need to check.

$$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{2}), \quad (\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, \frac{1}{2}), \quad (0, 0, 0)$$

We plug these values into our objective function, and find that $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$ are the solutions to our maximization problem. (Notice that the candidates with negative multipliers would be the solution to the minimization problem.)

18. Consider a standard utility maximization problem:

$$\max_{x_1, x_2} U(x_1, x_2)$$
$$s.t. \;\; p_1 x_1 + p_2 x_2 \leq I$$

where a consumer maximizes utility through two consumption goods, x_1, x_2, subject to their budget constraint. p_1, p_2 represent positive unit prices, and I is the agent's income. We will make the assumption that for each commodity bundle

(x_1, x_2),

$$\frac{\partial U}{\partial x_1} > 0 \quad or \quad \frac{\partial U}{\partial x_2} > 0$$

which is a common version of what's called a monotonicity assumption.

- Show that the tightness of the budget constraint (that the consumer spends all of their income purchasing goods) is a result of the monotonicity assumption.

Solution:

The usual constraint qualification is satisfied. We can now form the Lagrangian and write down the FONCs:

$$\begin{aligned}
&L = U(x_1, x_2) - \lambda(p_1 x_1 + p_2 x_2 - I) \\
&\frac{\partial L}{\partial x_1} = \frac{\partial U}{\partial x_1} - p_1 \lambda = 0 \\
&\frac{\partial L}{\partial x_2} = \frac{\partial U}{\partial x_2} - p_2 \lambda = 0 \\
&\lambda \geq 0 \\
&p_1 x_1 + p_2 x_2 \leq I \\
&\lambda(p_1 x_1 + p_2 x_2 - I) = 0
\end{aligned}$$

We have two cases to check. The first of which is when lambda is zero. However, if $\lambda = 0$, then both $\frac{\partial U}{\partial x_1} = 0$ and $\frac{\partial U}{\partial x_2} = 0$, a violation of our monotonicity condition.

Now let $\lambda > 0$. Then, due to the complementary slackness condition, we require that $p_1 x_1 + p_2 x_2 - I = 0$. Thus, the constraint is binding, and the consumer spends all of their income.

19. **Simon and Blume 18.11 Adapted** Consider the following minimization problem:

$$\begin{aligned}
&Minimize \;\; f(x, y) = 2y - x^2 \\
&subject \;\; to \;\; x^2 + y^2 \leq 1, \quad x \geq 0 \quad y \geq 0
\end{aligned}$$

- Check that the NDCQ are satisfied
- Write down the Lagrangian and FONCS
- What is the solution to the minimization problem?

<u>**Solution:**</u>

The Jacobian matrix of the constraints looks like

$$\begin{bmatrix} 2x & 2y \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

First notice that the gradient of $x^2 + y^2$ is only zero at the origin, far from the boundary of the constraint. Next, notice that not all three constraints can be binding. Given that we need to check only the rank of the Jacobian of binding constraints, its clear that the matrix will be of full rank for any combination of binding constraints. So the NDCQ will be satisfied at any solution candidate.

Before we write down the Lagrangian, we need to make sure that we write the constraints properly. Remember, this is now a minimization problem, so we want constraints to look like $g(x) \geq c$. So we will rewrite the first constraint as $-x^2 - y^2 \geq -1$.

$$L(x, y, \lambda_1, \lambda_2, \lambda_3) = 2y - x^2 - \lambda_1(-x^2 - y^2 + 1) - \lambda_2(x) - \lambda_3(y)$$

Now we can derive the FONCS:

$$\begin{aligned}
&\frac{\partial L}{\partial x} = -2x - \lambda_1(-2x) - \lambda_2 = 0 \\
&\frac{\partial L}{\partial y} = 2 - \lambda_1(-2y) - \lambda_3 = 0 \\
&\lambda_1(-x^2 - y^2 + 1) = 0, \quad \lambda_2(x) = 0, \quad \lambda_3(y) = 0 \\
&\lambda_1, \lambda_2, \lambda_3 = 0 \\
&x^2 + y^2 \leq 1, \quad x \geq 0 \;\; y \geq 0
\end{aligned}$$

Now, this may seem daunting given the number of equations and potential cases.

However, you can usually knock out quite a few cases by finding some requirement that a particular multiplier must be zero or positive. Let's rewrite the second equation.

$$2 + \lambda_1(2y) = \lambda_3$$

We know that both λ_1 and y are nonnegative, which requires that λ_3 be positive. Then, by the complementary slackness condition, we have y=0 and $\lambda_3 = 2$.

Now, we can check the cases where x is zero or positive. Let x=0. Then from the first equation, λ_2=0. From the complementary slackness condition, $\lambda_1 = 0$. Then, (0,0,0,0,2) is a candidate solution.

Let x>0. Then $\lambda_2 = 0$ from the complementary slackness condition, and $\lambda_1 = 1$ from the first order condition for x. If $\lambda_1 = 1$, then it must be the case that $x^2 + y^2 = 1$ from the complementary slackness condition, meaning that x=1 (we don't consider x=-1, as we have a nonnegativity constraint on x). So, (1,0,1,0,2) is a candidate solution as well. Now, because we are minimizing a continuous function over a compact set, by Weierstrass's Theorem we know the function will attain its global minimum. Because any constrained global min will be a local constrained min, then it must be one of the two identified candidates above. Plugging in our candidates, f(0,0)=0, while f(1,0)=-1. So (x,y)=(1,0) minimizes f(x,y) on the constraint set.