1 Problem Set 9 – Solutions

1.1 Problem 1

We are asked to minimize the function $d(x_1, y_1, x_2, y_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ subject to the constraints that $x_1 + y_1 = 10$ and $x_2^2 + 2y_2^2 = 1$. We will instead minimize the function $f(x_1, y_1, x_2, y_2) = d(x_1, y_1, x_2, y_2)^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2$. Clearly this is equivalent since the function $\phi(x) = x^2$ is monotonic on the non-negative reals, and clearly the image of d is a subset of the non-negative reals. Therefore we construct the function

$$F(x_1, y_1, x_2, y_2, \lambda_1, \lambda_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2 + \lambda_1(x_1 + y_1 - 10) + \lambda_2(x_2^2 + 2y_2^2 - 1)$$

and we are required to find all of the points such that $\frac{dF}{dx_i} = \frac{dF}{dy_i} = \frac{dF}{d\lambda_i} = 0$, for $i = 1, 2$. We solve as follows.

$$\frac{dF}{dx_1} = 2(x_1 - x_2) + \lambda_1 = 0 \Rightarrow 2(x_1 - x_2) = -\lambda_1$$

$$\frac{dF}{dy_1} = 2(y_1 - y_2) + \lambda_1 = 0 \Rightarrow 2(y_1 - y_2) = -\lambda_1$$

$$\Rightarrow x_1 - x_2 = y_1 - y_2$$

Note that clearly $x_1 - x_2 \neq 0$ since that would imply that there were a point of intersection between the two curves, but if we let $A = (a, b)$ be such a point of intersection then we see that $a = 10 - b \Rightarrow (10 - b)^2 + 2b^2 = 1$, but when we expand, and solve this quadratic using the quadratic formula, we get that there are no real roots, and hence there are no points of intersection between the two curves, so $x_1 \neq x_2$. Continuing,

$$\frac{dF}{dx_2} = 2(x_2 - x_1) + 2x_2\lambda_2 = 0 \Rightarrow 2x_2 = -\frac{x_2 - x_1}{\lambda_2}$$

$$\frac{dF}{dy_2} = 2(y_2 - y_1) + 4y_2\lambda_2 = 0 \Rightarrow 4y_2 = -\frac{y_2 - y_1}{\lambda_2}$$

But from above we get then that $x_2 = 2y_2$. So we get

$$1 = x_2^2 + 2y_2^2 = 6y_2^2 \Rightarrow y_2 = \pm\frac{1}{\sqrt{6}} \Rightarrow x_2 = \pm\frac{2}{\sqrt{6}}$$

This gives us two cases (one where $x_2, y_2 > 0$ and one where $x_2, y_2 < 0$). We consider each separately. Suppose first that $x_2 = \frac{2}{\sqrt{6}}, y_2 = \frac{1}{\sqrt{6}}$.

$\Rightarrow x_1 - \frac{2}{\sqrt{6}} = y_1 - \frac{1}{\sqrt{6}} \Rightarrow \frac{1}{\sqrt{6}} = x_1 - y_1 = 2x_1 - 10$

$\Rightarrow x_1 = 5 + \frac{1}{2\sqrt{6}} \Rightarrow y_2 = 5 - \frac{1}{2\sqrt{6}}$.
Now considering the other case, we have $x_2 = -\frac{2}{\sqrt{6}}, y_2 = -\frac{1}{\sqrt{6}}$
$\Rightarrow x_1 + \frac{2}{\sqrt{6}} = y_1 + \frac{1}{\sqrt{6}} \Rightarrow \frac{1}{\sqrt{6}} = y_1 - x_1 = 2y_1 - 10$
$\Rightarrow y_1 = 5 + \frac{1}{2\sqrt{6}}, x_1 = 5 - \frac{1}{2\sqrt{6}}$.
This gives us two sets of points which are local minima for the distance function. We simply calculate the distance function for each of them and compare them so see that the set of points which minimizes the distance between the two curves is
$\{P_1 = (5 + \frac{1}{2\sqrt{6}}, 5 - \frac{1}{2\sqrt{6}}), P_2 = (\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}})\}$.

1.2 problem 2

We first consider the interior of the ball,

$$B^o = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 < 1\}$$

Therefore, it suffices to find all local maxima and minima since clearly, if the maximum (or minimum) is not local then f does not achieve its max (or min) on B^o, but since B is compact, and all continuous functions attain their maximum (and minimum) on compact sets (and clearly f is compact), it must obtain its max (or min) on the boundary of B, $B - B^o$, which we will consider next.

Recall from class, that to find local maxima and minima we simply calculate partial derivatives and set them equal to zero. We get

$$\frac{df}{dx} = 3x^2 = 0 \Rightarrow x = 0$$

Similarly, $y = z = 0$. Therefore, the only critical point is $(0, 0, 0)$. Clearly this is not a maximum or minimum since for arbitrary $\epsilon > 0$ such that $\epsilon < \sqrt{\frac{1}{3}}$ we get that

$$f(-\epsilon, -\epsilon, -\epsilon) < f(0, 0, 0) < f(\epsilon, \epsilon, \epsilon)$$

Hence, the maximum and minimum must both be obtained on
$B - B^o = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$.
So now we define the function $g(x, y, z) = x^2 + y^2 + z^2 - 1$, and we may use Lagrange multipliers. We get the function
$h(x, y, z, \lambda) = f(x, y, z) + \lambda g(x, y, z)$. So we simply must find the points where $h_x = h_y = h_z = h_\lambda = 0$ ($h_x = \frac{dh}{dx}$, etc.).
We get,

$$h_x = 3x^2 + 2\lambda x = 0 \Rightarrow x = 0, -\frac{2\lambda}{3}$$

And we may solve similarly for y and z. Therefore, $x^2 = 0, \frac{4\lambda^2}{9}$, and similarly for y^2 and z^2. We now use the fact that $h_\lambda = g = 0$ to find λ and therefore, (x, y, z).
We now examine four cases, depending on the values of x, y, and z.

1.2.1 Case 1

$x = y = z = 0 \Rightarrow x^2 = y^2 = z^2 = 0$
$\Rightarrow h_\lambda = g(x, y, z) = 0 + 0 + 0 - 1 = -1 \neq 0 \Rightarrow\Leftarrow$.

1.2.2 Case 2

$x = -\frac{2\lambda}{3}, y = z = 0 \Rightarrow x^2 = \frac{4\lambda^2}{9}, y^2 = z^2 = 0$
$\Rightarrow h_\lambda = \frac{4\lambda^2}{9} - 1 = 0 \Rightarrow \lambda = \pm\frac{3}{2}$
$\Rightarrow x = \pm 1, y = z = 0 \Rightarrow f(x, y, z) = \pm 1$.

1.2.3 Case 3

$x = y = -\frac{2\lambda}{3}, z = 0 \Rightarrow x^2 = y^2 = \frac{4\lambda^2}{9}, z^2 = 0$
$\Rightarrow h_\lambda = \frac{4\lambda^2}{9} + \frac{4\lambda^2}{9} + 0 - 1 = 0 \Rightarrow \lambda = \pm\frac{3}{2\sqrt{2}}$
$\Rightarrow x = y = \pm\frac{1}{\sqrt{2}}, z = 0 \Rightarrow f(x, y, z) = \pm\frac{1}{\sqrt{2}}$.

1.2.4 Case 4

$x = y = z = -\frac{2\lambda}{3} \Rightarrow x^2 = y^2 = z^2 = \frac{4\lambda^2}{9}$
$\Rightarrow h_\lambda = \frac{4\lambda^2}{9} + \frac{4\lambda^2}{9} + \frac{4\lambda^2}{9} - 1 \Rightarrow \lambda = \pm\frac{\sqrt{3}}{2}$
$\Rightarrow x = y = z = \pm\frac{1}{\sqrt{3}} \Rightarrow f(x, y, z) = \pm\frac{1}{\sqrt{3}}$.
Now, clearly we see that

$$-1 < -\frac{1}{\sqrt{2}} < -\frac{1}{\sqrt{3}}, 1 > \frac{1}{\sqrt{2}} > \frac{1}{\sqrt{3}}$$

Therefore, the maximum and minimum of f are 1 and -1, on B. This occurs at the points $(1, 0, 0)$ and $(-1, 0, 0)$, respectively (or some reordering).

1.3 problem 3

We are asked to minimize the function $AM(x_1, \ldots, x_n) = \frac{x_1 + \cdots + x_n}{n}$ subject to the constraint that $x_1 \cdots x_n = 1, x_i \in \mathbb{R}$ for all i. So we get the function

$$F(x_1, \ldots, x_n, \lambda) = \frac{x_1 + \cdots + x_n}{n} + \lambda(x_1 \cdots x_n - 1)$$

and we take partial derivatives.

$\frac{dF}{dx_i} = \frac{1}{n} + \lambda x_1 \cdots x_{i-1}x_{i+1} \cdots x_n = \frac{1}{n} + \frac{\lambda x_i}{x_1 \cdots x_n} = \frac{1}{n} + \lambda x_i$.

So $\frac{dF}{dx_i} = 0 \Rightarrow x_i = \frac{-1}{n\lambda} = \alpha$ for all i.

So $x_i = \alpha$ for all i, and so $\alpha^n = 1$, but $\alpha \in \mathbb{R} \Rightarrow \alpha = \pm 1$ where $\alpha = 1$ if n is odd, and $\alpha = \pm 1$ if n is even. So if n is odd then our only choice is that $x_i = 1$ for all i, and if n is even then we get two options, either $x_i = 1$ for all i or else $x_i = -1$ for all i. Clearly the latter minimizes AM. And so we have minimized AM and shown that its minimum value subject to the given constraint is 1 if n is odd, and -1 if n is even (and we are allowed to consider points where $x_i < 0$).

Note The question has a typo in it, what he means to ask you to show is that the AM is greater than or equal to the GM for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that $x_i > 0$ for all i, (since otherwise this is clearly untrue, consider $x_1 = -2, x_2 = 0$.

It follows immediately from the above argument that the AM is greater than or equal to the GM for all $x_1, \ldots, x_n \in \mathbb{R}$ such that $x_1 \cdots x_n = 1$, since we minimized AM and saw that its minimum value is 1, and clearly if $x_1 \cdots x_n = 1$ then $GM = 1$, and so $AM \geq GM$, when $x_1 \cdots x_n = 1$. We will now show that it suffices to consider the case where $x_1 \cdots x_n = 1$.

First, notice that both AM and GM are homogeneouss (recall f homogeneous if $f(\lambda\vec{x}) = \lambda f(\vec{x})$ for all scalars λ). This is clear, simply by calculating, say $\vec{x} = (x_1, \ldots, x_n)$.

$$AM(\lambda\vec{x}) = \frac{\lambda x_1 + \cdots + \lambda x_n}{n} = \lambda\frac{x_1 + \cdots + x_n}{n} = \lambda AM(\vec{x})$$

$$GM(\lambda\vec{x}) = \sqrt[n]{\lambda x_1 \cdots \lambda x_n} = \sqrt[n]{\lambda^n x_1 \cdots x_n} = \lambda\sqrt[n]{x_1 \cdots x_n} = \lambda GM(\vec{x})$$

Now notice that for all $\vec{x} = (x_1, \ldots, x_n)$ such that $x_1, \ldots, x_n > 0$ there exists $\lambda \in \mathbb{R}, \lambda > 0$ such that $GM(\lambda\vec{x}) = 1$. Just set $\lambda = \frac{1}{GM(\vec{x})}$. But then we have that $AM(\lambda\vec{x}) \geq GM(\lambda\vec{x})$, and so since AM and GM are homogeneous, and $\lambda > 0$, we have that $AM(\vec{x}) \geq GM(\vec{x})$, as desired.

Note Because of the typo, I did not punish you if you didn't realize that this was what he wanted you to prove (also, some of you asked me what you were supposed to prove and I didn't answer with this, so whatev).

1.4 problem 4

Note All of these results are very intuitive, but many of them are notational nightmares.

1.4.1 (a)

It is given that $R_1, \ldots, R_m \subset R \subset \mathbb{R}^n$, for a rectangle R. Let $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, and let $R_i = [a_1^i, b_1^i] \times \cdots \times [a_n^i, b_n^i]$. If we assume that each rectangle is also an n dimensional rectangle, then we have that $a_k^i < b_k^i$ for all k, i. If each rectange needn't be in n dimensions, then the inequality is not strict. In either case, we have that $a_k \leq a_k^i \leq b_k^i \leq b_k$, for all k, i. So define the set $A_i = \{a_i, a_i^1, \ldots, a_i^m\}$, where if any a_i repeat, you only count it once. Similarly, define $B_i = \{b_i, b_i^1, \ldots, b_i^m\}$, and let $S_i = A_i \cup B_i$. We relabel the elements of S_i so that we have $S_i = \{s_1^i, \ldots, s_{n_i}^i\}$, where $s_1^i < s_2 < \cdots < s_{n_i}^i$, for all i (this is possible since we made sure that each s_j is distinct, and each is a real number, and there are finitely many, so clearly they can be ordered). Now, define $I_k^i = [s_k^i, s_{k+1}^i]$. Then I claim that the partition $P = \{I_{k_1}^1 \times \cdots \times I_{k_n}^n | k_i = 1, \ldots, n_i - 1\}$ is the desired partition. That is each R_i is the union of some rectangles in P, and all rectangles in P are disjoint. Both of these assertions are clear. Each rectangle must be disjoint since suppose not. Then there exists some $x \in \mathbb{R}^n$ such that x is contained in the interior of two distinct rectangles of P, say P_1 and P_2. Let $x = (x_1, \ldots, x_n)$, and let $P_1 = (s_{k_1}^1, s_{k_1+1}^1) \times \cdots \times (s_{k_n}^n, s_{k_n+1}^n)$, and $P_2 = (S_{K_1}^1, S_{K_1+1}^1) \times \cdots \times (S_{K_n}^n, S_{K_n+1}^n)$, where $s_j^i, S_J^i \in S_i$.

So $x \in P_1, P_2 \Rightarrow x_i \in (s_{k_i}^i, s_{k_i+1}^i), (S_{K_i}^i, S_{K_1+1}^i)$. But since all $s_j^i, S_J^i \in S_i$ if $k_i \neq K_i$ then WLOG, $s_{k_i+1}^i \leq S_{K_i}^i$, so the intervals are disoint. But then x_i cannot lie in both of them. Hence $k_i = K_i$ for all i, and so $P_1 = P_2$. So the rectangles of P are disjoint.

Note1 This method of creating this partition, P is a method that we will reuse throughout this problem. Let us call P the universal partition of $R^* = R_1 \cup \cdots \cup R_m$ (which is itself a partition, by definition).

Note2 I went right to the assumption that P_1 and P_2 were the interiors of rectangles, so the way I defined them they are not in fact rectangles, however they are surely the interiors of specific rectangles in the partition.

To show that any R_i is equal to some union of rectangles in P, we simply consider some $R_i = [s_{k_1}^1, s_{K_1}^1] \times \cdots \times [s_{k_n}^n, s_{K_n}^n]$. Then it is easy to see that the union of all rectangles in P all of whose sides (the intervals in $\mathbb{R}$ which make up the rectangle) are inside those of R_i, is equal to R_i. To see this, we show that the union is a subset of the rectangle, and that the rectangle is a subset of the union. The first is very easy to see. Simply choose some point, say x, in the union. If it's in the union, it must be in some individual rectangle, say $P^* \subset P$, which means that all of the coordinates of x are inside the corresponding interval of P^*, which is inside the corresponding interval of R_i, for all coordinates of x. Hence, $x \in R_i$. To show the other inclusion, let x now be in R_i, which means that each coordinate of x is inside the coresponding interval of R_i, which means that there exist $s_j, s'_j \in S_j$ such that $s_j \leq x_j \leq x'_j$, which means there exists $s^*_j \in S_j$ such that $s^*_j \leq x_j \leq s^{**}_j$ where s^*_j and s^{**}_j are consecutive elements of S_j. Therefore, x is in the rectangle of the partition which is given by each of these intervals for all of the $j = 1, \dots, n$, and so x is in the union of rectangles as desired, and we are done.

1.4.2 (b)

Clearly there exists some rectangle R such that $P \in R$. Then this result follows directly from (a).

1.4.3 (c)

Let us consider P^*_1 and P^*_2 to be the universal partitions of P_1 and P_2, respectively. Let S^i_j be the set of starting and ending points of P_j for dimension i. Then let $S^i = S^i_1 \cup S^i_2$. Then let P^* be the partition formed by considering all rectangles whose intervals are consecutive elements of S^i (consecutive elements can only be considered after S^i has been ordered, which is possible as there are a finite number of real elements). Then by exactly the same reasoning of (a), any rectangle in P_1 or P_2 can be formed from a disjoint union of rectangles in P^*, and we are done.

1.4.4 (d)

We first prove a lemma.

Lemma 1 *If we have a rectangle $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ and we partition it by cutting a single one of its intervals (say interval i) into two pieces giving us $R_1 = [a_1, b_1] \times \cdots \times [a_i, c] \times \cdots \times [a_n, b_n]$ and $R_2 = [a_1, b_1] \times \cdots \times [c, b_i] \times \cdots \times [a_n, b_n]$, then we have that $v(R) = v(R_1) + v(R_2)$.*

Proof Simply calculate. We get
$v(R_1) + v(R_2) =$
$= (b_1 - a_1) \cdots (c - a_i) \cdots (b_n - a_n) + (b_1 - a_1) \cdots (b_i - c) \cdots (b_n - a_n)$
$= [(b_1 - a_i) \cdots (b_{i-1} - a_{i-1})(b_{i+1} - a_{i+1}) \cdots (b_n - a_n)][b_i - c + c - a_i]$
$= (b_1 - a_1) \cdots (b_i - a_i) \cdots (b_n - a_n) = v(R)$, as desired.

Therefore, by repeated application of this theorem, it follows that the volume of any rectangle is equal to the sum of the rectangles in any partition of that rectangle, which is exactly what we wanted to show, and so we are done.

1.4.5 (e)

We use (d) and we see that if P_1 and P_2 are two partitions of P, then

$$\sum_{R \in P_1} v(R) = v(P) = \sum_{R \in P_2} v(R)$$

and we are done.

1.4.6 (f)

First we prove a useful lemma.

Lemma 2 *If R_1 and R_2 are rectangles, then $v(R_1 \cup R_2) + v(R_1 \cap R_2) = v(R_1) + v(R_2)$.*

Proof First note that if two rectangles are not disjoint, then their intersection is a rectangle (this is just because if two intervals intersect, their intersection is an interval). Therefore, we may partition R_1 and R_2 each into two different rectangles. We partition R_1 into M_1 and N and we partition R_2 into M_2 and N, where $N = R_1 \cap R_2$ and $M_i = R_i - N$ (notice that if the triangles are disjoint, then $N = \emptyset$ and so we would be partitioning R_i into R_i and $\emptyset$). In this way we have constructed a partition of $R_1 \cup R_2$ into the disjoint rectangles M_1, M_2, N. Now we use (e), and we get that
$v(R_1 \cup R_2) + v(R_1 \cap R_2) = v(M_1) + v(N) + v(M_2) + v(N) = v(R_1) + v(R_2)$, as desired, and our lemma is complete.

Now, we first note that if we have P and Q polygons with $P \subset Q$, then we may write Q as $P \cup R$ where $R \cap P = \emptyset$ (note that if $Q = P$ then $R = \emptyset$). But then from our lemma we have that
$v(Q) = v(P \cup R) = v(P) + v(R) - v(P \cap R) = v(P) + v(R)$, and clearly $v(A) \geq 0$ for all polygons A (with equality if and only if $A = \emptyset$).
Hence, $v(Q) \geq v(P)$ as desired.

1.4.7 (g)

Suppose we have a polygon $P = R_1 \cup \cdots \cup R_n$ for R_i rectangles. Let $D_2 = \{x \in P | x \in R_i \cap R_j, i \neq j\}$.
Similarly, define $D_k = \{x \in P | x \in R_{m_1} \cap \cdots \cap R_{m_k}, i \neq j \Rightarrow m_i \neq m_j\}$.
So first notice that D_i is a polygon for all i, as the intersection of two or more rectangles is a rectangle, which makes D_i a colection of rectangles, or a polygon. Secondly, notice that
$D_n \subset D_{n-1} \subset \cdots \subset D_2 \Rightarrow v(D_n) \leq \cdots \leq v(D_2)$ (from (f)). Also, notice that
$v(R_1)+\cdots+v(R_n)-v(D_2)+v(D_3)-\cdots+(-1)^{n+1}v(D_n) = v(R_1\cup\cdots\cup R_n)$.
This is because if you are measureing the volume of the union of several rectangles, it suffices to count all of the individual volumes of the rectangles, then subtract out the volumes of the points which you have counted twice, but then you have to add back in the volumes of all of the points you removed twice, namely all of the points in three rectangles, etc. This fact is also very provable using induction.

But now we are done, since if n is eventhen we end by subtracting some non-negative quantity. So if we can show it for n odd (where we end by adding a non-negative quantity) we can definitely show it for n even. So suppose n is odd.
$v(R_1 \cup \cdots \cup R_n) =$
$= v(R_1) + \cdots + v(R_n) - (v(D_2) - v(D_3)) - \cdots - (v(D_{n-1} - v(D_n))$
$\leq v(R_1) + \cdots + v(R_n)$, since $v(D_{2k} \geq D_{2k+1}$, as shown above. But this is what we wanted to prove and so we are done.