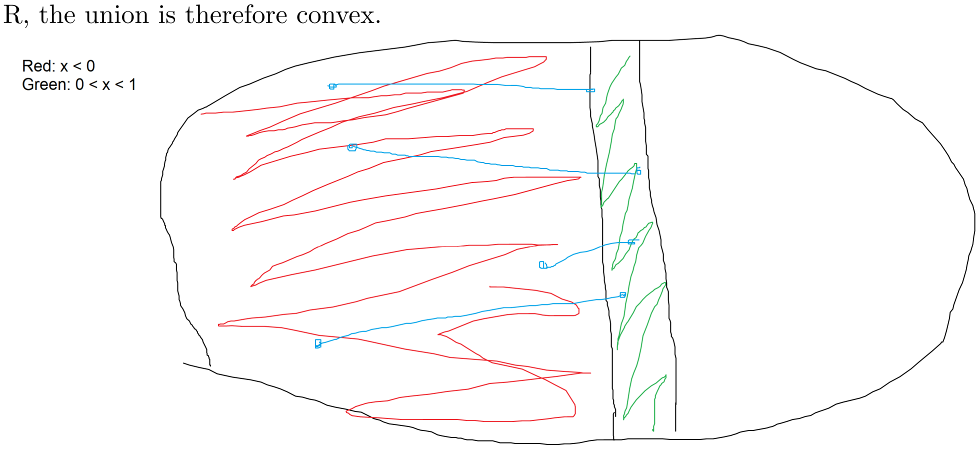
1 Problem 1

1.1 a

We can say that for all $x_1...x_M$, $||x - x_0||_2 <= ||x - x_k||_2$ for $k = 1...M$. This can be rewritten as $||x - x_0||_2 - ||x - x_k||_2 <= 0$ for $k = 1...M$. This means that for each $k = 1...M$, we can say there is a half space that exists for each k. The total number of points that is contained in this set is the intersection of all these half spaces. Since each half space is convex, this set is the intersection of M convex spaces, meaning it too is convex since intersection preserves convexity.

1.2 b

As can be seen graphically, the union of these two sets still connects in a straight line between the two spaces, and since it never leaves the boundary of the shape of R, the union is therefore convex.

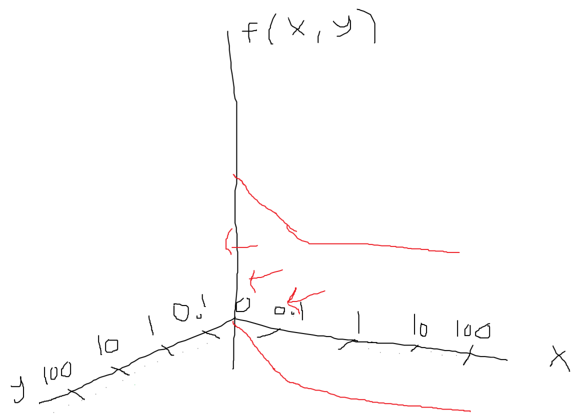


1.3 c

Since x is positive semidefinite, we know that $1^T x 1 >= 0$, or that $1^T x >= 0$. If we also know that $1^T x <= 1$, then this means that all x in this subset are representative of a convex hull, as we saw in lecture. This is therefore convex.

1.4 d

We can see that the shape generated from this function f(x, y) is represented as a halfspace. This space runs along the line y=1/x. However, since the line y=1/x is inherently curved, this means that we cannot draw straight lines from every point for $y_1 >= 1/x_1$ to another point $y_2 >= 1/x_2$. This means that this is not a convex set as you cannot draw straight lines between the curved points without going underneath the line $y = 1/x$.

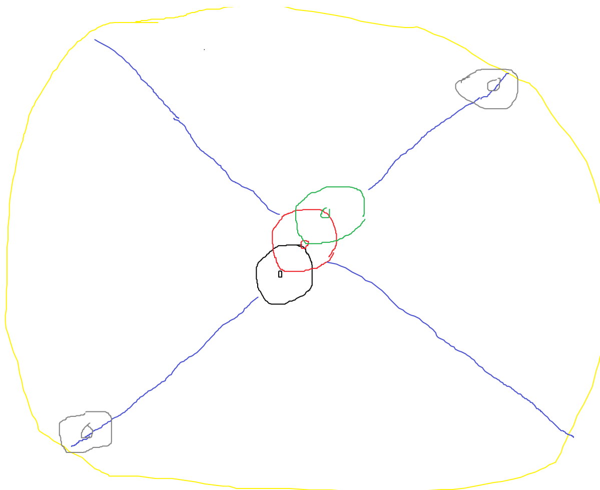


1.5 e

If A is a positive semidefinite matrix, we can say that every vector x in R^n will make it so that $x^T Ax >= 0$. Therefore, this set of x is every vector in the space R^n, which means that this set is convex because the space R^n is convex by nature.

1.6 f

As can be seen in the graphic, in R^n, each vector x_i has a n-dimensional sphere for radius 1 around it that represents all the vectors that are part of the set. In the end, this shape represents a closed, n-dimensional sphere itself, meaning that every vector from x_0 to x_{n-1} within it can be connected by a straight line segment that never exits the bounds of the largest, n-dimensional sphere. Therefore, this is a convex set.



2 Problem 2

For a cone in two dimensions, we can describe the formula as $\sqrt{Ax^2 + Bt^2} <= c$, for some constant c that is maximum possible value in the set C. This means $\sqrt{Ax_1^2 + Bt_1^2} <= c$ and $\sqrt{Ax_2^2 + Bt_2^2} <= c$. For some $0 <= \theta <= 1$, since θ is guaranteed to shrink the magnitude of any number it is multiplied by, we can say that $\theta\sqrt{Ax^2 + Bt^2} <= \theta c <= c$ and $(1-\theta)\sqrt{Ax^2 + Bt^2} <= (1-\theta)c <= c$. If we combine these two, we can say that $\theta\sqrt{Ax^2 + Bt^2} + (1-\theta)\sqrt{Ax^2 + Bt^2} <= \theta c + (1-\theta)c$. We know that $\theta + (1-\theta) = 1$, so this is equal to $\theta\sqrt{Ax^2 + Bt^2} + (1-\theta)\sqrt{Ax^2 + Bt^2} <= c$. Therefore, the set containing $\theta[x_1, t_1] + (1-\theta)[x_2, t_2]$ is within the set C because it is still less than the maximum possible value c, meaning it is convex.

3 Problem 3

For the set S_{-a}, we can say that the set only contains elements x_i, where for S_{-a}, $||x_i - x|| <= a$. We know that the Euclidean ball of radius r is a convex set, which is the set of all elements x centered around x_c, $||x - x_c|| <= r$. We can see the direct parallel between these two shapes, with radii, centers, and points around the center. If we say that we have the set X_a that is the Euclidean ball with radius r equal to a, then we can say that S_{-a} is the intersection of X_a and S. Since both of these are convex sets, and the intersection of convex sets is convex, then we know that S_{-a} is convex.

4 Problem 4

4.1 a

$f(x_1, x_2) = x_1 x_2 on R^2_{++}$. We know that x_1, x_2 are both greater than 0 and real because of the nature of R^2_{++}. The Hessian of this function is equal to

$$\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}$$

This matrix is not positive semidefinite because we know it has eigenvalues -1 and 1, and -1 is a negative eigenvalue. Therefore, this is not a convex function.

It is also not concave, however, because if we take $-f(x_1, x_2) = -x_1 x_2$, the Hessian of this is has the exact same eigenvalues as the Hessian of f, so this function is neither concave nor convex.

4.2 b

$f(x_1, x_2) = x_1/x_2 on R^2_{++}$. We know that x_1, x_2 are both greater than 0 and real because of the nature of R^2_{++}. The Hessian of this function is equal to

$$\begin{matrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2^3 \end{matrix}$$

This tells us that the function is not convex. For example, if $x_1 = 100$ and $x_2 = 5$, then this will yield a Hessian with a negative eigenvalue.

The following is the Hessian of $-f(x_1, x_2)$ $\begin{matrix} 0 & 1/x_2^2 \\ 1/x_2^2 & -2x_1/x_2^3 \end{matrix}$

This tells us that the function is not concave either. For example, if $x_1 = 5$ and $x_2 = 100$, then this will yield a Hessian with a negative eigenvalue.

5 Problem 5

5.1 a

The Hessian of this function is equal to the 2x2 matrix: $\begin{matrix} 2 & -a \\ -a & 2 \end{matrix}$

This means that in order for the function f to be convex, we need this to be positive semidefinite, or when a is in-between -2 and 2.

5.2 b

The first derivative of this function is equal to $2(A+\alpha I)x$. The derivative of this function (or the second derivative of the original function) is equal to $2(A+\alpha I)$. This means that if we need this to be positive semidefinite, α must be positive and greater than any negative diagonal values of A or negative and smaller than the positive eigenvalues, while A has non-negative values in the non-diagonal entries.

6 Problem 6

6.1 a

s=1 is minimized by v=0, where $log(1) = 0 = -1 + 1$. s=2 is minimized by v=-log(2), where $log(2) = 0.69 = -1 + 2 * e^{-0.69} + 1 =$ 0.69. This patterns continues for all s that are greater than 0. We can even see the same for $0 < s < 1$, such as s=0.5 where $log(0.5) = -0.69 = -1 + 0.5(e^{0.69} - 0.69 \ = 0.69$. Therefore, for any $s > 0$, $log(s) = -1 + (se^v - v), wherev = -log(s)$.

6.2 b

If we take $f(z) = log(\sum_{i=1}^{n} e^{z_i + v} - v)$, then since we know that $log(s) = -1 + (se^v - v), wherev = -log(s)$, we can say that $f(z) = -1 + ((\sum_{i=1}^{n} e^{z_i})e^v - v) = -1 + ((\sum_{i=1}^{n} e^{z_i} + v) - v), wherev = -log(\sum_{i=1}^{n} e^{z_i})$. In other words, we can say this equals $-1 + ((\sum_{i=1}^{n} e^{z_i} + v) - v), wherevminimizesthesubexpressionintheequation.$

6.3 c

The concept of partial minimization says that if f(z) is convex, then the minimum element of the set is convex as well. Since we have the minimum value of this set being equal to $-1 + ((\sum_{i=1}^{n} e^{z_i} + v) - v), wherevminimizesthesubexpressionintheequation$, if we can prove this is convex then we know that the set is convex. We know that this element is convex because for every z_i in the vector z, the elements will be in the shape of a half space. Therefore, this element is convex, meaning the set of f(z) is also convex.

7 Problem 7

7.1 a

7.2 i

If the step size is too small, the algorithm runs repeatedly before hitting the maximum number of iterations. Once that happens, the resulting line is incredibly linear, meaning the result is not appropriately found.

7.3 ii

If the step size is too big, the algorithm runs repeatedly before hitting the maximum number of iterations. Once that happens, the resulting line is incredibly parabolic, meaning the result is not appropriately found as it is overestimated.

7.4 iii

Step size of 0.5 converges in 1 iteration.

7.5 iv

The Lipschitz constant β is defined as $||\delta f(x) - \delta f(y)||/||x - y||$. For this notebook, that is equal to $||2x - 2y||/||x - y||$, which makes $\beta = 2$. we set the step size to $1/\beta$, we get a graph that only needs to run for one iteration. When we set step size equal to $1/\beta + -0.1$, we get a graph that is very close to linear from top to bottom, with a slight exponential curve. It only needs to run for 8 iterations in this case.

7.6 b

7.7 i

For a step size of 0.75, if $x_0 = 0.1$, then we see that the curve is a steep curve from 0.01 down with 57 iterations.

If $x_0 = 1$, then we see that the curve is a steep curve from 1 down with 68 iterations.

If $x_0 = 5$, then we see that the curve is a steep curve from 25 down with 75 iterations.

If $x_0 = 50$, then we see that the curve is a steep curve from 2500 down with 86 iterations.

This shows us that a higher guess will result in more iterations, and a similar shaped curve.

7.8 ii

For c = -2, we see that we hit the maximum number of iterations. For c = -1, we have 128 iterations. For c = 0, we have 66 iterations. For c = 1, we have 128 iterations. For c = 2, we see that we have 30 iterations. The closer c is to 2, we see that the lines more closely resemble a perfect circle. The farther we get, the more they look like ellipses until they become parallel lines at c=2 and c=-2.