Exercise 1. Consider the space $X = C^k([0,1])$ of k times continuously differentiable functions. We often impose boundary conditions so we take the subset

$$D = \left\{ \varphi \in X : \varphi(0) = 2, \frac{\partial}{\partial x}\varphi(1) = 3 \right\} \subseteq X$$

of functions with a prescribed function value on the left, and prescribed derivative on the right end of the interval. (This clearly only makes sense if $k \geqslant 1$.)

i) Show that this is an affine subspace of X.

ii) Show that D is convex.

iii) State what the tangent space D' of D is.

Answer

To show that D is affine, we must show it is of the form $v + D_0$ where D_0 is indeed a subspace of X. To that effect, consider the subspace

$$D_0 = \left\{ \varphi \in X : \varphi(0) = 0, \frac{\partial}{\partial x}\varphi(1) = 0 \right\}.$$

For $\varphi, \psi \in D_0$ we have

$$c\varphi(0) + \psi(0) = c\cdot 0 + 0 = 0 \quad \text{and} \quad \frac{\partial}{\partial x}(c\varphi(x) + \psi(x))\,|_{\,x=1} = c\varphi_x(1) + \psi_x(1) = 0,$$

so it is indeed a subspace. Now if we consider $v = 3x + 2$, then when exploring the conditions imposed on our functions we have that for $\varphi \in D_0$

$$3(0) + 2 + \varphi(0) = 2 \quad \text{and} \quad \frac{\partial}{\partial x}(3x + 2 + \varphi(x))\,|_{\,x=1} = 3 + 0 = 3.$$

So it holds that $D = (3x + 2) + D_0$ and therefore, it is an affine subspace.
Consider now two functions $\varphi, \psi \in D$, we are to show that any convex combination of them is in D. To that effect note that for any $t \in [0,1]$ we have

$$t\psi(0) + (1-t)vf(0) = 2t + 2(1-t) = 2t + 2 - 2t = 2.$$

When differentiating we get

$$\frac{\partial}{\partial x}\left(t\psi(x) + (1-t)\varphi(x)\right)\,|_{\,x=1} = t\psi_x(1) + (1-t)\varphi_x(1) = 3t + 3 - 3t = 3.$$

In conclusion, convex combinations of functions in our space are still in our space. It follows that D is convex.
Finally, reacall that the tangent space of D at a point is

$$D'(x_0) = \{\, v \in X : \; v + x_0 \in D \,\}$$

and as D is a linear affine subspace, D' is independent of the selection of x_0. Regardless, we claim that $D' = D_0$. Observe that if we add any element of D_0 to an element of D we don't switch the boundary conditions. This means that $D_0 \subseteq D'$, and if we take such a $v \in D'$, then there is a $\varphi \in D$ such that $v + x_0 = \varphi$. This means that $v = \varphi - x_0$ so when considering the boundary conditions we see that v's boundary conditions are precisely those of D_0. It follows that the tangent space is precisely D_0. *This makes sense, as D is a linear space, the tangent space must be its slopes. But that's itself. However, D' must be a subspace, so that's why it's D_0.*

Exercise 2. Take the function

$$f(x, y) = \min\{\, |x|, |y| \,\}\, \mathrm{sgn}(x).$$

In class, we talked about the fact that the directional (Gateaux) derivative satisfies

$$Df((x, y); v) = \langle \nabla f(x, y) | v \rangle\,.$$

For the current function, at the origin, we have $\nabla f(0) = 0$[1], but the directional derivative is not zero for all v.

i) Let's build intuition. Plot this function. Then use this visualization to convince yourself (and the reader of your answer) graphically that the statement about $\nabla f(0) = 0$ is true, as well as that the statement that the directional derivative is not zero in other directions is also true.

ii) Explain the discrepancy. What does this imply for the viability of the idea that we can look for points with $\nabla f = 0$ when searching for minima of functions?

[1] That is because the gradient is defined as $\nabla f = (\partial_x f, \partial_y f)^\intercal$ and because the function f is constant (equal to zero) along both the x and y axes

Answer

Imagine first the function $z = |x|$. When viewing the slice at $y = 0$ we get a $\bigvee$ shaped curve. So as there's no dependence on the y variable, the shape of the surface $z = |x|$ is a $\bigvee$ shaped trough follwing the y axis. The surface $z = |y|$ has the same kind of story but rotated 90°. If we reduce the opacity of one of the troughs and increase it for the other, we can see that the section of the dim trough is covered above by the more visible one. This can't happen if we are taking the minimum, so we must eliminate that part.
If we cut the shape by a plane $z = c$, then the resulting surface is an $\times$ shaped trough. The final detail is considering what happens we multiply by $\operatorname{sgn}(x)$. On the right half of plane we still have our trough cut in half so it is $\vdash$ shaped now. However on the left, we take the same mirror image of our trough but flip it 180° with respect to the yz plane.