HW 4 SOLUTIONS

Due on Feb 19, 2020 (in class).

The transpose of a matrix A will be written A^t. The standard inner product in $\mathbf{R}^n$ will be denoted $\langle\ ,\ \rangle$.

In the solutions below, as in the lectures, we distinguish between an n-tuple and a row vector. For us $(x_1, \ldots, x_n)$ is the transpose of $[\, x_1\ x_2\ \cdots\ x_n \,]$. In other words $(x_1, x_2, \ldots, x_n)$ is a typographically convenient way of writing a column vector.

Velocity vector of a path and the gradient of a function. Let U be an open subset of $\mathbf{R}^n$. Let $f\colon U \to \mathbf{R}$ be a $\mathscr{C}^1$ function. The *gradient of* f is the map

$$\boldsymbol{\nabla} f\colon U \to \mathbf{R}^n$$

given by

$$\boldsymbol{\nabla} f(\boldsymbol{x}) = (D_1 f(\boldsymbol{x}), \ldots, D_n f(\boldsymbol{x})) = f'(\boldsymbol{x})^t.$$

If $\boldsymbol{\gamma}\colon (a,b) \to U$ is a smooth path, i.e. $\boldsymbol{\gamma}$ is a $\mathscr{C}^1$ map, then the *velocity of* $\boldsymbol{\gamma}$ at a point $\theta \in (a,b)$ is

$$\boldsymbol{v}_{\boldsymbol{\gamma}}(\theta) = \boldsymbol{\gamma}'(\theta) \in \mathbf{R}^n.$$

If $\boldsymbol{p} = \boldsymbol{\gamma}(\theta)$, we often write $\boldsymbol{v}_{\boldsymbol{\gamma}}(\boldsymbol{p})$ for $\boldsymbol{v}_{\boldsymbol{\gamma}}(\theta)$. The default assumption when this notation is used is that $\boldsymbol{\gamma}$ does not "visit" $\boldsymbol{p}$ more than once (easily achieved by restricting the domain of $\boldsymbol{\gamma}$). Sometimes the map $\boldsymbol{\gamma}'\colon (a,b) \to \mathbf{R}^n$ is also called the velocity vector, even though, rightly speaking, it is a family of velocity vectors. We might indulge in these two types of looseness occassionaly for ease of writing. We say $\boldsymbol{\gamma}$ *passes through* $\boldsymbol{p} \in U$ if there is a $\theta \in (a,b)$ such that $\boldsymbol{\gamma}(\theta) = \boldsymbol{p}$.

(1) Let f and $\boldsymbol{\gamma}$ be as above. Show that

$$\frac{\mathrm{d}f(\boldsymbol{\gamma}(t))}{\mathrm{d}t} = \langle \boldsymbol{v}_{\boldsymbol{\gamma}}(t),\, \boldsymbol{\nabla} f(\boldsymbol{\gamma}(t))\rangle$$

for $t \in (a,b)$.

Solution: By the chain rule

$$\begin{aligned}\frac{\mathrm{d}f(\boldsymbol{\gamma}(t))}{\mathrm{d}t} = f'(\boldsymbol{\gamma}(t))\boldsymbol{\gamma}'(t) &= f'(\boldsymbol{\gamma}(t))'\boldsymbol{v}_{\boldsymbol{\gamma}}(t)\\ &= \boldsymbol{\nabla} f(t)^t \boldsymbol{v}_{\boldsymbol{\gamma}}(t)\\ &= \langle \boldsymbol{\nabla} f(t), \boldsymbol{v}_{\boldsymbol{\gamma}}(t)\rangle.\end{aligned}$$

$\square$

(2) Let f and $\boldsymbol{\gamma}$ be as above. Suppose S is the hypersurface in U given by the equation $f(\boldsymbol{x}) = c$ where c is a constant. Suppose the path $\boldsymbol{\gamma}$ is actually a path in S, i.e. $\boldsymbol{\gamma}(t) \in S$ for all $t \in (a,b)$. Show that $\boldsymbol{v}_{\boldsymbol{\gamma}}(t)$ is orthogonal to $\boldsymbol{\nabla} f(\boldsymbol{\gamma}(t))$ for all $t \in (a,b)$.

Solution: Clearly $(f \circ \boldsymbol{\gamma})(t) \equiv c$, whence $\frac{\mathrm{d}f(\boldsymbol{\gamma}(t))}{\mathrm{d}t} = 0$ for all $t \in (a,b)$. By Problem (1), this means $\langle \boldsymbol{\nabla} f(t), \boldsymbol{v}_{\boldsymbol{\gamma}}(t)\rangle = 0$ for $t \in (a.b)$ and this is what we were asked to prove. $\square$

Tangent spaces. Let U be open in $\mathbf{R}^n$ and $\boldsymbol{g}\colon U \to \mathbf{R}^m$ a $\mathscr{C}^1$ map, $\boldsymbol{c} \in \mathbf{R}^m$ a point in the image of $\boldsymbol{g}$ and M the subset of U given by the equation $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{c}$. Suppose that $\boldsymbol{p} \in M$ is a point such that the rank of $\boldsymbol{g}'(\boldsymbol{p})$ is m. Let $V_{\boldsymbol{p}}$ be the null space of $\boldsymbol{g}'(\boldsymbol{p})$. The space $T_{\boldsymbol{p}} = V_{\boldsymbol{p}} + \boldsymbol{p}$ is called the tangent space to M at $\boldsymbol{p}$.

(3) Show that $V_{\boldsymbol{p}}$ is the space spanned by all velocity vectors $\boldsymbol{v}_{\boldsymbol{\gamma}}(\boldsymbol{p})$ for $\mathscr{C}^1$ paths $\boldsymbol{\gamma}$ taking values in M and passing through $\boldsymbol{p}$. [**Hint:** It is easy to see that the velocity vectors at $\boldsymbol{p}$ of paths in M passing through $\boldsymbol{p}$ lie in $V_{\boldsymbol{p}}$. To show that the space spanned by them is all of $V_{\boldsymbol{p}}$ use the Implicit Function Theorem.]

Solution: Let $\boldsymbol{g} = (g_1, \dots, g_m)$. The rows of $\boldsymbol{g}'$ are the transposes of the gradients $\boldsymbol{\nabla} g_i$. More precisely, the i^{th} row of $\boldsymbol{g}'$ is $(\boldsymbol{\nabla} g_i)^t$, $i = 1, \dots, m$. Now, from the previous problem we know that $\langle \boldsymbol{\nabla} g_i, \boldsymbol{v}_{\boldsymbol{\gamma}}(\boldsymbol{p})\rangle = 0$, $i = 1, \dots, m$, for any $\mathscr{C}^1$ path $\boldsymbol{\gamma}$ taking values in M and passing through $\boldsymbol{p}$ (once). This means that the rows of $\boldsymbol{g}'(\boldsymbol{p})$ when multiplied on the left to $\boldsymbol{v}_{\boldsymbol{\gamma}}(\boldsymbol{p})$, yield 0, whence $\boldsymbol{g}'(\boldsymbol{p})\boldsymbol{v}_{\boldsymbol{\gamma}}(\boldsymbol{p}) = 0$. Thus, by definition, $\boldsymbol{v}_{\boldsymbol{\gamma}}(\boldsymbol{p}) \in V_{\boldsymbol{p}}$. Thus if H is the vector subspace of $\mathbf{R}^n$ spanned by velocity vectors at $\boldsymbol{p}$, then $H \subset V_{\boldsymbol{p}}$.

Since rank $\boldsymbol{g}'(\boldsymbol{p}) = m$, the Jacobian matrix $(J\boldsymbol{g})(\boldsymbol{p})$ must have m linearly independent columns. Recall that $(J\boldsymbol{g})(\boldsymbol{p})$ is a $m \times n$ matrix and its j^{th} column (for $j \in \{1, \dots, n\}$) is $(D_j\boldsymbol{g})(\boldsymbol{p})$. Without loss of generality (by re-ordering our co-ordinates) we may assume $(D_{d+1}\boldsymbol{g})(\boldsymbol{p}), \dots, (D_n\boldsymbol{g})(\boldsymbol{p})$ are linearly independent. Let

$$\boldsymbol{p} = (\boldsymbol{a}, \boldsymbol{b})$$

with $\boldsymbol{a} \in \mathbf{R}^d$ and $\boldsymbol{b} \in \mathbf{R}^m$. By the implicit function theorem, we have an open connected neighbourhood W of $\boldsymbol{a}$ in $\mathbf{R}^d$, and a unique $\mathscr{C}^1$ map $\boldsymbol{f}\colon W \to \mathbf{R}^m$ such that $\boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{b}$, and $(\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x})) \in M$ for every $\boldsymbol{x} \in W$.

Let $\boldsymbol{e}_1, \dots, \boldsymbol{e}_d$ be the standard basis of $\mathbf{R}^d$ and for $i \in \{1, \dots, d\}$ let

$$\boldsymbol{\sigma}_i(t) = \boldsymbol{a} + t\boldsymbol{e}_i.$$

We can find $\eta > 0$ such that $\boldsymbol{\sigma}_i(-\eta, \eta) \subset W$ for all $i = 1, \dots, d$. The right way to think about $\boldsymbol{\sigma}_i$ is as the path passing through $\boldsymbol{a}$ at time $t = 0$ along the i^{th} coordinate axis, at a speed of one distance unit per time unit. In fact the velocity vector of $\boldsymbol{\sigma}_i$ at $\boldsymbol{a}$ is $\boldsymbol{e}_i$. Let $\boldsymbol{\gamma}_i\colon (-\eta, \eta) \to \mathbf{R}^n$ be the map

$$\boldsymbol{\gamma}_i(t) = (\boldsymbol{\sigma}_i(t),\ \boldsymbol{f}(\boldsymbol{\sigma}_i(t))), \qquad (i = 1, \dots, d).$$

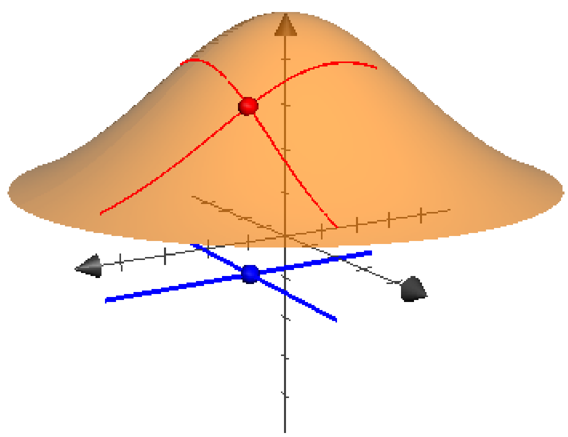
Note that for each $i = 1, \dots, d$, $\boldsymbol{\gamma}_i$ takes values in M since $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ is the solution of the implicit equation $\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{c}$. Then the velocity vector $\boldsymbol{v}_i := \boldsymbol{v}_{\boldsymbol{\gamma}_i}(\boldsymbol{p})$ is given by

$$\boldsymbol{v}_i = \left(\boldsymbol{e_i}, \tfrac{\mathrm{d}(\boldsymbol{f} \circ \boldsymbol{\sigma}_i)}{\mathrm{d}t}\big|_{t=0}\right) = \left(\boldsymbol{e_i}, \boldsymbol{f}'(\boldsymbol{a})\tfrac{\mathrm{d}(\boldsymbol{\sigma}_i)}{\mathrm{d}t}\big|_{t=0}\right) = (\boldsymbol{e}_i, \boldsymbol{f}'(\boldsymbol{a})\boldsymbol{e}_i).$$

It is clear that $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$ are linearly independent since $\boldsymbol{e}_1, \dots, \boldsymbol{e}_d$ are. In greater detail, if $\alpha_1, \dots, \alpha_d \in \mathbf{R}$ are such that if $\sum_{i=1}^d \alpha_i \boldsymbol{v}_i = \boldsymbol{0}$ then $\sum_{i=1}^d \alpha_i(\boldsymbol{e}_i, \boldsymbol{f}'(\boldsymbol{a})\boldsymbol{e}_i) = \boldsymbol{0}$, whence $\sum_{i=1}^d \alpha_i \boldsymbol{e}_i = \boldsymbol{0}$. This means $\alpha_i = 0$ for every i, i.e. $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d$ are linearly independent. (The picture given after the remark below may help.) Now, by definition of H, $\boldsymbol{v}_1, \dots, \boldsymbol{v}_d \in H \subset V_{\boldsymbol{p}}$. Since the rank of $\boldsymbol{g}'(\boldsymbol{a}) = m$, the dimension of its null space $V_{\boldsymbol{p}}$ is $n - m = d$, i.e. $\dim V_{\boldsymbol{p}} = d$. Thus $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d$ is a basis of $V_{\boldsymbol{p}}$ and we are done. $\square$

Remark: We have actually proved that every vector in $V_{\boldsymbol{p}}$ at $\boldsymbol{p}$ is a velocity vector at $\boldsymbol{p}$ of a $\mathscr{C}^1$ path in M. With the notations of the solution, we know $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d$ is a basis for $V_{\boldsymbol{p}}$. Suppose $\boldsymbol{v} \in V_{\boldsymbol{p}}$. Then $\boldsymbol{v} = \sum_{i=1}^{d} \alpha_i \boldsymbol{v}_i$ for a unique sequence of scalars $\alpha_1, \ldots, \alpha_d$. Let $\boldsymbol{\sigma}(t) = \boldsymbol{a} + t \sum_{i=1}^{d} \alpha_i \boldsymbol{e}_i$. Since $\boldsymbol{\sigma}(0) = \boldsymbol{a} \in W$, therefore $\boldsymbol{\sigma}$ takes values in W in an open interval containing 0, say $(-\delta, \delta)$, and let us restrict the domain of $\boldsymbol{\sigma}$ to this interval. Let $\boldsymbol{\gamma} = (\boldsymbol{\sigma}, \boldsymbol{f} \circ \boldsymbol{\sigma})$. Then $\boldsymbol{\gamma}(t) \in M$ for $t \in (-\delta, \delta)$ and $\boldsymbol{\gamma}(0) = \boldsymbol{p}$. An easy calculation shows that the velocity of $\boldsymbol{\gamma}$ at 0 is $\boldsymbol{v}$, proving the assertion at the beginning of this remark.

In the following picture, the surface is M, the red dot is $\boldsymbol{p}$, and the blue dot $\boldsymbol{a}$. The blue lines "on the floor" are the curves $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ passing through $\boldsymbol{a}$ and the red curves on the surface are their "lifts", i.e. they are $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ and they pass through $\boldsymbol{p}$. The velocity vectors of $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_1$ at $t = 0$ form a basis for $V_{\boldsymbol{p}}$.



(4) Let W be an open set in $\mathbf{R}^n$ and $\boldsymbol{f} \colon W \to \mathbf{R}^m$ a $\mathscr{C}^1$ function. Let Γ be the graph of $\boldsymbol{f}$, i.e. $\Gamma = \{(\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{x})) \in W \times \mathbf{R}^m \mid \boldsymbol{x} \in W\} \subset \mathbf{R}^{m+n}$. Let $\boldsymbol{p} \in W$, and set $A = \boldsymbol{f}'(\boldsymbol{p})$ and $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{p})$. Show that the tangent space to Γ at $\boldsymbol{q} = (\boldsymbol{p}, \boldsymbol{b}) \in \Gamma$ makes sense (in other words, see that the conditions given for the definition to make sense are satisfied, especially the condition on the rank of a derivative), and show that $T_{\boldsymbol{q}}$ is given by the system of linear equations in $\mathbf{R}^{m+n} = \mathbf{R}^n \times \mathbf{R}^m$:

$$\boldsymbol{y} = A(\boldsymbol{x} - \boldsymbol{p}) + \boldsymbol{b}.$$

Solution: Let $\boldsymbol{\varphi} \colon W \times \mathbf{R}^m \to \mathbf{R}^m$ be the map given by the formula

$$\boldsymbol{\varphi}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y} \qquad (\boldsymbol{x} \in W, \boldsymbol{y} \in \mathbf{R}^m).$$

It is clear that Γ is given by the equation $\boldsymbol{\varphi}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$ on $W \times \mathbf{R}^m \subset \mathbf{R}^{n+m}$. A little thought shows that the block matrix form of the Jacobian matrix of $\boldsymbol{\varphi}$ at $\boldsymbol{q}$ is

$$(J\boldsymbol{\varphi})(\boldsymbol{q}) = \big[(J\boldsymbol{f})(\boldsymbol{p}) \quad | \quad -I\big]$$

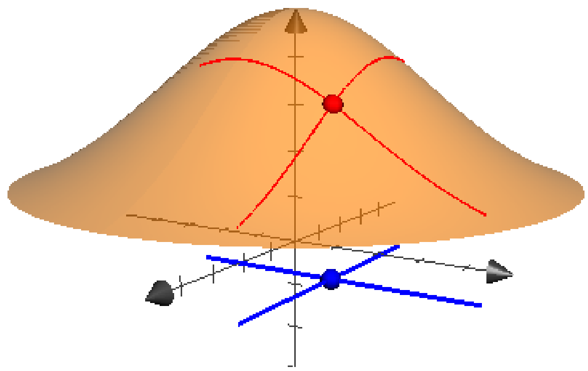
where I is the $m \times m$ identity matrix. It follows that $\boldsymbol{\varphi}'(\boldsymbol{q})$ has rank m. By definition, $V_{\boldsymbol{q}}$ is the null space of $\boldsymbol{\varphi}'(\boldsymbol{q})$, i.e. the null space of $(J\boldsymbol{\varphi})(\boldsymbol{q})$. Since

$T_{\boldsymbol{q}} = V_{\boldsymbol{q}} + \boldsymbol{q}$, and since $\boldsymbol{q} = (\boldsymbol{p}, \boldsymbol{b})$, $(\boldsymbol{x}, \boldsymbol{y}) \in T_{\boldsymbol{q}}$ if and only if $(\boldsymbol{x} - \boldsymbol{p}, \boldsymbol{y} - \boldsymbol{b}) \in V_{\boldsymbol{q}}$. Equivalently

$$\begin{bmatrix} (J\boldsymbol{f})(\boldsymbol{p}) & -I \end{bmatrix} \begin{bmatrix} \boldsymbol{x} - \boldsymbol{p} \\ \boldsymbol{y} - \boldsymbol{b} \end{bmatrix} = \boldsymbol{0}.$$

This is equivalent to $(J\boldsymbol{f})(\boldsymbol{p})(\boldsymbol{x} - \boldsymbol{p}) - (\boldsymbol{y} - \boldsymbol{b}) = \boldsymbol{0}$. Since $(J\boldsymbol{f})(\boldsymbol{p})$ is the matrix representation of $\boldsymbol{f}'(\boldsymbol{p}) = A$ in the standard bases of $\mathbf{R}^n$ and $\mathbf{R}^m$, this gives $A(\boldsymbol{x} - \boldsymbol{p}) - (\boldsymbol{y} - \boldsymbol{b}) = \boldsymbol{0}$, as required. □

Let me fill in space with some more pictures associated with problem 3.



Here is a view "from the below the ground", i.e. from below the xy-plane. Actually from below five units from the xy-plane, so from quite a bit in the basement.