1 Recap

- Proof of chain rule.
- Level sets and tangent planes.

2 Vector fields

Recall that a vector field is a function $\vec{F} : S \subset \mathbb{R}^n \to \mathbb{R}^m$. Recall that $\vec{F}$ is said to be continuous at $\vec{a}$ given $\epsilon > 0$ there exists a $\delta > 0$ such that whenever $|\vec{r} - \vec{a}| < \delta$, then $|\vec{F}(\vec{r}) - \vec{F}(\vec{a})| < \epsilon$. $\vec{F}$ is continuous if and only if its component scalar fields are so.
Let $\vec{a} \in S$ be an interior point. $\vec{F}$ is said to have a derivative $\nabla_{\vec{v}}\vec{F}(\vec{a})$ along $\vec{v}$ at the point $\vec{a}$ if $\nabla_{\vec{v}}\vec{F}(\vec{a}) = \frac{\vec{F}(\vec{a}+h\vec{v})-\vec{F}(\vec{a})}{h}$ exists. It is easy to prove that the directional derivative exists if and only if the directional derivative of each component exists. Moreover, $\nabla_{\vec{v}}\vec{F}(\vec{a}) = (\nabla_{\vec{v}}F_1(\vec{a}), \nabla_{\vec{v}}F_2(\vec{a}), \ldots)$ (HW). In particular, one can talk of partial derivatives of $\vec{F}$. $\vec{F}$ is said to be *differentiable* at $\vec{a}$ if there exists a linear map $D\vec{F}_{\vec{a}} : \mathbb{R}^n \to \mathbb{R}^m$ such that $\lim_{\vec{h}\to\vec{0}} \frac{\|\vec{F}(\vec{a}+\vec{h})-\vec{F}(\vec{a})-D\vec{F}_{\vec{a}}(\vec{h})\|}{\|\vec{h}\|} = 0$. The map $D\vec{F}_{\vec{a}}$ is called the derivative or total derivative of $\vec{F}$ at $\vec{a}$.
(HW) $\vec{F}$ is differentiable if and only if each component is so. Moreover, $D\vec{F}(\vec{v}) = \nabla_{\vec{v}}\vec{F}$.
In other words, $D\vec{F}(\vec{v}) = \begin{bmatrix} \nabla F_1^T \\ \nabla F_2^T \\ \vdots \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \end{bmatrix}$.
For example, if $\vec{E} = (y, -x)$, then since each component is a polynomial, it is differentiable and $\nabla E_1 = (0,1)$, $\nabla E_2 = (-1, 0)$. Thus $D\vec{E} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.
Another example, if $\vec{F}(r,\theta) = (r\cos(\theta), r\sin(\theta))$, then $D\vec{F} = \begin{bmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{bmatrix}$

Theorem: If $\vec{F}$ is differentiable at $\vec{a}$, then it is continuous at $\vec{a}$.
Proof: $\|\vec{F}(\vec{a}+\vec{h}) - \vec{F}(\vec{a}) - D\vec{F}_{\vec{a}}(\vec{h})\| < \|\vec{h}\|\frac{\epsilon}{2}$ whenever $0 < \|\vec{h}\| < \delta < 1$. By the triangle inequality $\|\vec{F}(\vec{a}+\vec{h}) - \vec{F}(\vec{a})\| < \|D\vec{F}_{\vec{a}}(\vec{h})\| + \frac{\epsilon}{2}$.
At this juncture, we prove a useful linear algebraic lemma: Suppose A is an $m \times n$ matrix and $\vec{v} \in \mathbb{R}^n$. Then $\|A\vec{v}\| \leq C_A\|\vec{v}\|$ where $C_A = \sum_i \|A_i\|$ (where A_i is the i^{th} row of A). Proof of lemma: $\|A\vec{v}\| = \|\sum_i \langle A_i, \vec{v}\rangle\| \leq \sum_i |\langle A_i, \vec{v}\rangle e_i| \leq \sum_i \|A_i\|\|\vec{v}\|$ by Cauchy-Schwarz.
Returning back to the proof of the theorem, $\|\vec{F}(\vec{a}+\vec{h}) - \vec{F}(\vec{a})\| < \|\vec{h}\|C_{D\vec{F}_{\vec{a}}} + \frac{\epsilon}{2} < \epsilon$ if $\|\vec{h}\|$ is small enough. □

3 Chain rule in greater generality

Recall that we wanted to know that if $T(x,y)$ and $\tilde{T}(x(r,\theta), y(r,\theta))$ are functions, then what is $\tilde{T}_r$ in terms of T_x, T_y etc. Roughly speaking, $\tilde{T}(x(r+h, \theta+k), y(r+h, \theta+k)) \approx \tilde{T}(x(r) + h\frac{\partial x}{\partial r} + \frac{\partial x}{\partial \theta}k, y(r) + h\frac{\partial y}{\partial r} + k\frac{\partial y}{\partial \theta}$ which is further approximately equal to

$\tilde{T}(x(r,\theta),y(r,\theta))+\frac{\partial \tilde{T}}{\partial x}\left(h\frac{\partial x}{\partial r}+\frac{\partial x}{\partial \theta}k\right)+\frac{\partial \tilde{T}}{\partial y}\left(h\frac{\partial y}{\partial r}+\frac{\partial y}{\partial \theta}k\right)$. In terms of matrices, it is $(\nabla T)\left[\begin{array}{cc} x_r=\cos(\theta) & x_\theta = \\ y_r=\sin(\theta) & y_\theta = \end{array}\right.$
The statement of the chain rule in this case is: If $\vec{g}(u,v)=(x(u,v),y(u,v))$ is differentiable at (a,b) and $f(x,y)$ is differentiable at $\vec{g}(a,b)$, then $h(u,v)=f\circ\vec{g}(u,v)=f(x(u,v),y(u,v))$ is differentiable at (a,b) and $\nabla h=\nabla f D\vec{g}$.
Suppose we have $\vec{F}(x,y)=(F_1(x,y),F_2(x,y))$ and $\vec{g}(u,v)=(x(u,v),y(u,v))$, then what must the derivative of $\vec{H}=\vec{F}\circ\vec{g}$ at $\vec{a}$ look like? Going by the Chain rule stated earlier, it ought to be $\begin{bmatrix}\nabla H_1\\ \nabla H_2\end{bmatrix}=\begin{bmatrix}\nabla F_1 D\vec{g}\\ \nabla F_2 D\vec{g}\end{bmatrix}=D\vec{F}_{\vec{g}(\vec{a})}D\vec{g}_{\vec{a}}$. In other words, we expect the derivative linear map to be a *composition* of the maps or the matrix to be a product of derivative matrices.
Theorem: Let $\vec{G}:S\subset\mathbb{R}^n\to\mathbb{R}^m$ be a vector field differentiable at an interior point $\vec{a}\in S$. Let $\vec{F}:U\subset\mathbb{R}^m\to\mathbb{R}^p$ be a vector field defined on U containing $\vec{G}(S)$. Suppose $\vec{g}(\vec{a})$ is an interior point of U and $\vec{F}$ is differentiable at $\vec{g}(\vec{a})$. Then $\vec{H}=\vec{F}\circ\vec{G}:S\subset\mathbb{R}^n\to\mathbb{R}^p$ is differentiable at $\vec{a}$ and $D\vec{H}_{\vec{a}}=D\vec{F}_{\vec{g}(\vec{a})}\circ D\vec{G}_{\vec{a}}$ or in terms of matrices, it is the product of matrices.
Proof: Let's prove it for scalar fields. Applying it to each component of a vector field is good enough. Now $f(\vec{g}(\vec{a})+\vec{y})-f(\vec{g}(a))-Df_{\vec{g}(\vec{a})}[\vec{y}]=\|\vec{y}\|E(\vec{g}(\vec{a}),\vec{y})$ where E goes to 0 as $\vec{y}$ goes to $\vec{0}$ by assumption of differentiability of f. Likewise, $\vec{g}(\vec{a}+\vec{h})=\vec{g}(\vec{a})+D\vec{g}_{\vec{a}}[\vec{h}]+\|\vec{h}\|\vec{F}(\vec{a},\vec{h})$ where $\vec{F}$ goes to 0 as $\vec{h}\to\vec{0}$. Let $\vec{y}=\vec{g}(\vec{a}+\vec{h})-\vec{g}(\vec{a})$. It goes to $\vec{0}$ as $\vec{h}$ does (why?). Thus, $f(\vec{g}(\vec{a}+\vec{h}))-f(\vec{g}(\vec{a}))-Df_{\vec{g}(\vec{a})}D\vec{g}_{\vec{a}}[\vec{h}]=Df_{\vec{g}(\vec{a})}[\vec{y}]+\|\vec{y}\|E(\vec{g}(\vec{a}),\vec{y})-Df_{\vec{g}(\vec{a})}D\vec{g}_{\vec{a}}[\vec{h}]$ which equals $\|\vec{h}\|Df_{\vec{g}(\vec{a})}\vec{F}(\vec{a},\vec{h})+\|\vec{y}\|E(\vec{g}(\vec{a}),\vec{y})=\|\vec{h}\|H(\vec{a},\vec{h})$ where H goes to 0 (why?) $\square$