MATH 425b ASSIGNMENT 8 SOLUTIONS
SPRING 2016
Prof. Alexander

Chapter 9:

(30)(a) We have $\mathbf{p}'(t) = \mathbf{x}$ so by the Chain Rule,

$$(*) \qquad h^{(1)}(t) = h'(t) = f'(\mathbf{p}(t))\mathbf{p}'(t) = (D_1 f(\mathbf{p}(t)), \dots, D_n f(\mathbf{p}(t)))\cdot\mathbf{x} = \sum (D_{i_1} f)(\mathbf{p}(t)) x_{i_1},$$

where the sum is over all $i_1 \le n$. This proves the result for the first derivative, $k = 1$. Suppose $2 \le k \le m$ and the result is true for the $(k-1)$st derivative. Then

$$h^{(k)}(t) = \frac{d}{dt} h^{(k-1)}(t) = \frac{d}{dt} \sum (D_{i_2 \dots i_k} f)(\mathbf{p}(t)) x_{i_2} \cdots x_{i_k},$$

where the sum is over all $(k-1)$-tuples $(i_2, \dots, i_k)$. Since this is a finite sum, we can differentiate term by term, and using the Chain Rule we get as in $(*)$ above

$$\begin{aligned} \frac{d}{dt} \sum_{(i_2,\dots,i_k)} (D_{i_2 \dots i_k} f)(\mathbf{p}(t)) x_{i_2} \cdots x_{i_k} &= \sum_{(i_2,\dots,i_k)} \left[\sum_{i_1} (D_{i_1} D_{i_2 \dots i_k} f)(\mathbf{p}(t)) x_{i_1} \right] x_{i_2} \cdots x_{i_k} \\ &= \sum_{(i_1,\dots,i_k)} (D_{i_1 \dots i_k} f)(\mathbf{p}(t)) x_{i_1} x_{i_2} \cdots x_{i_k}. \end{aligned}$$

This shows the result is true for the kth derivative, and by induction it is then true for all $1 \le k \le m$.

(b) Define $r(\mathbf{x})$ by

$$f(\mathbf{a} + \mathbf{x}) = \sum_{k=0}^{m-1} \frac{1}{k!} \sum (D_{i_1 \cdots i_k} f)(\mathbf{a}) x_{i_1} \cdots x_{i_k} + r(\mathbf{x}).$$

Since $f(\mathbf{a} + \mathbf{x}) = h(1)$, by part (a) this is the same as

$$h(1) - \sum_{k=0}^{m-1} \frac{h^{(k)}(0)}{k!} = r(\mathbf{x}).$$

By Taylor's theorem 5.15 and part (a), then, we can express $r(\mathbf{x})$ as

$$r(\mathbf{x}) = \frac{h^{(m)}(t)}{k!} = \sum_{(i_1,\dots,i_m)} (D_{i_1 \dots i_m} f)(\mathbf{p}(t)) x_{i_1} x_{i_2} \cdots x_{i_m} \quad \text{for some } t \in (0,1).$$

Since the function $\sum_{(i_1,\ldots,i_m)} |D_{i_1\ldots i_m} f|$ is continuous, there exists a finite M and a $\delta > 0$ such that

$$|\mathbf{x}| < \delta \implies |\mathbf{p}(t) - \mathbf{a}| < \delta \text{ for all } t \in (0,1) \implies \sum_{(i_1,\ldots,i_m)} |D_{i_1\ldots i_m} f(\mathbf{p}(t))| \leq M \text{ for all } t \in (0,1)$$

$$\implies |r(\mathbf{x})| \leq \sum_{(i_1,\ldots,i_m)} |D_{i_1\ldots i_m} f(\mathbf{p}(t))| \cdot |\mathbf{x}|^m \leq M|\mathbf{x}|^m = o(|\mathbf{x}|^{m-1}) \text{ as } \mathbf{x} \to 0.$$

(I)(a) $\gamma'(z) = (D_2 u(y_0, z), 0, 1)$ so $\gamma'(z_0) = (D_2 u(y_0, z_0), 0, 1)$. Similarly $\tilde{\gamma}'(z_0) = (D_1 u(y_0 z_0), 1, 0)$.

(b) Suppose φ, has a local maximum on S at (x_0, y_0, z_0). Then since it takes values in S, $\varphi(\gamma(z))$ has a local maximum (equal to $\varphi(x_0, y_0, z_0)$) at $z = z_0$, and $\varphi(\tilde{\gamma}(y))$ has a local maximum at $y = y_0$. Therefore by the chain rule,

$$0 = \frac{d}{dz}\varphi(\gamma(z))\big|_{z=z_0} = \varphi'(\gamma(z_0)) \cdot \gamma'(z_0) = \nabla\varphi(x_0, y_0, z_0) \cdot \gamma'(z_0),$$

so $\nabla\varphi(x_0, y_0, z_0) \perp \gamma'(z_0)$. Similarly

$$0 = \frac{d}{dy}\varphi(\tilde{\gamma}(y))\big|_{y=y_0} = \varphi'(\tilde{\gamma}(y_0)) \cdot \tilde{\gamma}'(y_0) = \nabla\varphi(x_0, y_0, z_0) \cdot \tilde{\gamma}'(y_0),$$

so $\nabla\varphi(x_0, y_0, z_0) \perp \tilde{\gamma}'(y_0)$. Since $\nabla\varphi(x_0, y_0, z_0)$ is perpendicular to the two vectors that span the translated tangent plane, it is perpendicular to the entire plane.

f itself is constant on S, everywhere equal to 0, so it has a local maximum at every point of S, in particular at (x_0, y_0, z_0). Therefore by the preceding, $\nabla f(x_0, y_0, z_0)$ is perpendicular to the translated tangent plane. Since $\nabla\varphi(x_0, y_0, z_0)$ perpendicular to the same plane in $\mathbb{R}^3$, and $\nabla f(x_0, y_0, z_0) \neq 0$ (because we assumed $D_1 f(x_0, y_0, z_0) \neq 0$), $\nabla\varphi(x_0, y_0, z_0)$ must be a scalar multiple of $\nabla f(x_0, y_0, z_0)$.

(II)(a) The quadratic part is $D_{11}f(\mathbf{a})x_1^2 + D_{12}f(\mathbf{a})x_1x_2 + D_{21}f(\mathbf{a})x_1x_2 + D_{22}f(\mathbf{a})x_2^2$, which is the same as $\mathbf{x}^t H\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} D_{11}f(\mathbf{a}) & D_{12}f(\mathbf{a}) \\ D_{21}f(\mathbf{a}) & D_{22}f(\mathbf{a}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ if we set $H = \begin{bmatrix} D_{11}f(\mathbf{a}) & D_{12}f(\mathbf{a}) \\ D_{21}f(\mathbf{a}) & D_{22}f(\mathbf{a}) \end{bmatrix}$. (This is the *Hessian matrix.*) By Theorem 9.41, this is a symmetric matrix.

(b) Since H is symmetric, we have $H = U^t DU$ for some unitary U and diagonal $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, where the λ_i are eigenvalues of H. Therefore the quadratic part of the Taylor polynomial satisfies

$$\mathbf{x}^t H\mathbf{x} = \mathbf{x}^t U^t DU\mathbf{x} = (U\mathbf{x})^t D(U\mathbf{x}) = \lambda_1(U\mathbf{x})_1^2 + \lambda_2(U\mathbf{x})_2^2 \geq \lambda_2((U\mathbf{x})_1^2 + (U\mathbf{x})_2^2) = \lambda_2|\mathbf{x}|^2.$$

(c) From exercise 30, we have

$$f(\mathbf{a} + \mathbf{x}) = f(\mathbf{a}) + D_1 f(\mathbf{a})x_1 + D_2 f(\mathbf{a})x_2 + \mathbf{x}^t H\mathbf{x} + r(|\mathbf{x}|),$$

with $r(\mathbf{x}) = o(|\mathbf{x}|^2)$ as $x \to 0$. Since by assumption $Df(\mathbf{a}) = 0$, we have $D_1 f(\mathbf{a}) = D_2 f(\mathbf{a}) = 0$, so in fact this simplifies to

$$f(\mathbf{a} + \mathbf{x}) = f(\mathbf{a}) + \mathbf{x}^t H \mathbf{x} + r(|\mathbf{x}|).$$

Since $r(\mathbf{x}) = o(|\mathbf{x}|^2)$, there exists $\delta > 0$ such that $0 < |\mathbf{x}| < \delta$ implies $|r(\mathbf{x})| \leq \lambda_2 |\mathbf{x}|^2/2$, which in turn implies

$$f(\mathbf{a} + \mathbf{x}) \geq f(\mathbf{a}) + \lambda_2 |\mathbf{x}|^2 - \frac{1}{2}\lambda_2 |\mathbf{x}|^2 > f(\mathbf{a}).$$

This shows f has a local minimum at $\mathbf{a}$.