# Math 241 Section F1H, Fall 2007

# Differentiability on $n$-dimensional spaces

## Assumptions and notations:

- $\mathbf{R}^n$ denotes the $n$-dimensional Euclidean space.
- Bold-faced letters denote elements of such a space, and subscripts are used to indicate components: e.g., $\mathbf{x} = (x_1, \ldots, x_n)$.
- $|\mathbf{x}|$ denotes the Euclidean norm of $\mathbf{x}$: $|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \cdots}$.
- Unless otherwise stated, $\mathbf{f}$ denotes a function from $\mathbf{R}^n$ to $\mathbf{R}^m$, and its component functions are denoted by $f_1, f_2, \ldots$: $\mathbf{f} = (f_1, \ldots, f_m)$.

## Differentiability:

Given a function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ and a point $\mathbf{a}$ in $\mathbf{R}^n$, $\mathbf{f}$ is said to be **differentiable at a** if there exists an $m \times n$ matrix $\mathbf{A}$ such that

$$\lim_{|\mathbf{h}| \to 0} \frac{|\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) - \mathbf{A} \cdot \mathbf{h}|}{|\mathbf{h}|} = 0, \tag{1}$$

where $\mathbf{A} \cdot \mathbf{h}$ denotes the matrix product between $\mathbf{A}$ (an $m \times n$ matrix) and $\mathbf{h}$ (an $n$-dimensional column vector, regarded as an $n \times 1$ matrix). The matrix $\mathbf{A}$ is called the **derivative matrix of f at a** and is denoted by $\mathbf{A} = \mathbf{Df}(\mathbf{a})$.

## Taylor's formula:

If $\mathbf{f}$ is differentiable at $\mathbf{a}$, then

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) = \mathbf{P}_1(\mathbf{h}) + \mathbf{R}_1(\mathbf{h}),$$

where

$$\mathbf{P}_1(\mathbf{h}) = \mathbf{f}(\mathbf{a}) + \mathbf{Df}(\mathbf{a}) \cdot \mathbf{h},$$

is the called the **first order Taylor polynomial for f at a** and serves as a first order approximation to $\mathbf{f}(\mathbf{a}+\mathbf{h})$, and $\mathbf{R}_1(\mathbf{h})$ is a remainder term satisfying

$$\lim_{|\mathbf{h}| \to 0} \frac{|\mathbf{R}_1(\mathbf{h})|}{|\mathbf{h}|} = 0.$$

(The latter formula is simply a rewritten version of (1), using the notation $\mathbf{R}_1$.) This is analogous to the definition of Taylor polynomials and remainder terms for ordinary one-variable functions; see, e.g., formula (10), p. 709, in Edwards/Penney.

## Explicit form of derivative matrix:

If $\mathbf{f} = (f_1, \ldots, f_m)$ is differentiable at some point $\mathbf{a}$, then the partial derivatives $\partial f_i / \partial x_j$, $i = 1, 2, \ldots, m$ $j = 1, 2, \ldots, n$, all exist at this point, and they make up the derivative matrix $\mathbf{Df}$ at that point; i.e., we have (with the argument $\mathbf{a}$ left out for simplicity of notation):

$$\mathbf{Df} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

## Special cases:

- **Ordinary one-variable derivative:** For functions from $\mathbf{R}^1$ to $\mathbf{R}^1$, $\mathbf{D}f(a)$ is the usual derivative $f'(a)$, regarded as a $1 \times 1$ matrix.
- **Derivatives of vector functions:** A vector function $\vec{r}(t)$ is a function from $\mathbf{R}^1$ to $\mathbf{R}^m$, where $m$ is the dimension of the space in which the vectors "live", so the derivative matrix $\mathbf{D}\vec{r}$ for such functions is an $m \times 1$ matrix. This matrix turns out to be the componentwise derivative $\vec{r}\,'$, regarded as a column vector.
- **Gradients:** If $f$ is a **scalar** function of several variables, i.e., a function from $\mathbf{R}^n$ to $\mathbf{R}^1$, then the derivative matrix $\mathbf{Df}$ is the gradient of $f$, $\nabla f$, regarded as a row vector, or an $1 \times n$ matrix.
- **Jacobians:** For a coordinate transformation $\mathbf{T}(u, v) = (x(u, v), y(u, v))$ in two dimensional space (e.g., the polar coordinate transformation $\mathbf{T}(r, \theta) = (x, y) = (r\cos\theta, r\sin\theta)$), the derivative matrix is the Jacobian matrix of this transformation: $\mathbf{DT} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$. (This case came up in some HW problems from Section 12.7 of Edwards/Penney.)

## Matrix form of chain rule:

Given two functions $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ and $\mathbf{g} : \mathbf{R}^m \to \mathbf{R}^p$, their **composition** $\mathbf{h} = \mathbf{g} \circ \mathbf{f}$ is the function from $\mathbf{R}^n$ to $\mathbf{R}^p$ defined by $\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$. Using the above derivative concept, the chain rule takes a simple and elegant form: If $\mathbf{f}(\mathbf{x})$ is differentiable (in the above sense) at $\mathbf{x} = \mathbf{a}$ with derivative $\mathbf{Df}(\mathbf{a})$, and $\mathbf{g}$ is differentiable at $\mathbf{y} = \mathbf{f}(\mathbf{a})$, then $\mathbf{h}$ is differentiable at $\mathbf{x} = \mathbf{a}$ with derivative

$$\mathbf{Dh}(\mathbf{a}) = \mathbf{Dg}(\mathbf{f}(\mathbf{a})) \cdot \mathbf{Df}(\mathbf{a}).$$

Written out in component form, this means (with the arguments $\mathbf{a}$ and $\mathbf{f}(\mathbf{a})$ left out for simplicity of notation):

$$\begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_p}{\partial x_1} & \cdots & \frac{\partial h_p}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial g_p}{\partial y_1} & \cdots & \frac{\partial g_p}{\partial y_m} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$