Lecture 5: Section 2.3 Cont.. If $\mathbf{F} : \mathbf{R}^3 \to \mathbf{R}^3$ is given by $\mathbf{F}(\mathbf{x}) = \big(f_1(x_1, x_2, x_3), f_2(x_1, x_2, x_3), f_3($
then the **derivative matrix** of partial derivatives is defined by

$$\mathbf{DF} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial x_1} & \frac{\partial \mathbf{F}}{\partial x_2} & \frac{\partial \mathbf{F}}{\partial x_3} \end{bmatrix}$$

Here the second expression mean that the we think of $\mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$ as a column vector and $\frac{\partial \mathbf{F}}{\partial x_1}$ as the derivative of the column vector. Similarly we define the $m \times n$ matrix $\mathbf{DF}$ for a function $\mathbf{F} : \mathbf{R}^n \to \mathbf{R}^n$.

A special case is the **gradient** of a function $f : \mathbf{R}^n \to \mathbf{R}$ given by

$$\mathbf{grad}\, f = \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

A function $\mathbf{F} : \mathbf{R}^n \to \mathbf{R}^m$ is called differentiable at $\mathbf{a}$ if

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{a}) - \mathbf{DF}(\mathbf{a})(\mathbf{x} - \mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|} = 0$$

Here $\mathbf{DF}(\mathbf{a})(\mathbf{x} - \mathbf{a})$ is matrix multiplication of the $m \times n$ matrix or column vector $\mathbf{DF}(\mathbf{a})$ by the $n \times 1$ matrix $\mathbf{x} - \mathbf{a}$. If $f : \mathbf{R}^2 \to \mathbf{R}$ then

$$\mathbf{D}f(\mathbf{a})(\mathbf{x}-\mathbf{a}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) & \frac{\partial f}{\partial x_2}(\mathbf{a}) \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ x_2 - a_2 \end{bmatrix} = \frac{\partial f}{\partial x_1}(\mathbf{a})(x_1 - a_1) + \frac{\partial f}{\partial x_2}(\mathbf{a})(x_2 - a_2)$$

If $\mathbf{F} : \mathbf{R}^n \to \mathbf{R}^m$ then

$$\mathbf{DF}(\mathbf{a})(\mathbf{x} - \mathbf{a}) = \begin{bmatrix} \frac{\partial \mathbf{F}}{\partial x_1}(\mathbf{a}) & \dots & \frac{\partial \mathbf{F}}{\partial x_n}(\mathbf{a}) \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ \dots \\ x_n - a_n \end{bmatrix}$$

$$= \frac{\partial \mathbf{F}}{\partial x_1}(\mathbf{a})(x_1 - a_1) + \dots + \frac{\partial \mathbf{F}}{\partial x_n}(\mathbf{a})(x_n - a_n) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \\ \vdots \\ \frac{\partial f_m}{\partial x_1} \end{bmatrix} (x_1 - a_1) + \dots + \begin{bmatrix} \frac{\partial f_1}{\partial x_n} \\ \vdots \\ \frac{\partial f_m}{\partial x_n} \end{bmatrix} (x_n - a_n)$$

We can think of the derivative of $\mathbf{F}$ at the point $\mathbf{a} \in \mathbf{R}^n$ as the linear map $\mathbf{DF} : \mathbf{R}^n \to \mathbf{R}^m$, mapping the vector $\mathbf{h} = (h_1, ..., h_n)$ to the directional derivative

$$\mathbf{DF}(\mathbf{a})\mathbf{h} = \lim_{t \to 0} \frac{\mathbf{F}(\mathbf{a} + t\mathbf{h}) - \mathbf{F}(\mathbf{a})}{t} = \frac{\partial \mathbf{F}}{\partial x_1}(\mathbf{a})h_1 + ... + \frac{\partial \mathbf{F}}{\partial x_n}(\mathbf{a})h_n,$$

Ex If A is an $m \times n$ matrix we define $\mathbf{F}(\mathbf{x}) = A\mathbf{x} : \mathbf{R}^n \to \mathbf{R}^m$. Then $\mathbf{DF} = A$.

Th If f is differentiable then it is continuous.
Pf See book.

Th If the partial derivatives are continuous in a neighborhood that the function is differentiable in the neighborhood.
Pf See book.

Section 2.4.
Th $\partial_x \partial_y f(x, y) = \partial_y \partial_x f(x, y)$
Pf See book.

We also went over Newton's method for finding an approximation of a solution to $f(x) = 0$.