Math 350 Exam 1 Topics

- Vectors in $\mathbb{R}^n$, length/distance, dot product:

$$\|\mathbf{x}\| = \left(x_1^2 + \cdots + x_n^2\right)^{1/2}, \qquad \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$$

- Functions $f : \mathbb{R}^n \to \mathbb{R}^m$, composition
- Limits (ϵ-δ definition, but no proofs), continuity:

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{c} \text{ means}$$
$$\text{for all } \epsilon > 0, \text{ there exists } \delta > 0 \text{ such that}$$
$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|f(\mathbf{x}) - \mathbf{c}\| < \epsilon$$

  $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous if $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$.

- Partial derivatives. If $f : \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} = (x_1, \ldots, x_n) \mapsto (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$,

$$\frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_n) = \lim_{h \to 0} \frac{f_i(x_1, \ldots, x_j + h, \ldots, x_n) - f_i(x_1, \ldots, x_n)}{h}$$

- Total derivative. Definition: $\mathbf{D}f(\mathbf{x}_0)$ is the unique linear function $\mathbf{D}f(\mathbf{x}_0) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{D}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$$

  As you proved on PS3, the matrix for $\mathbf{D}f(\mathbf{x}_0)$ is the matrix of partial derivatives:

$$\mathbf{D}f(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{bmatrix}$$

- Special cases of total derivative:
  - Velocity: for a path/curve $\mathbf{c} : \mathbb{R} \to \mathbb{R}^n$, $t \mapsto (c_1(t), \ldots, c_n(t))$,

$$\mathbf{c}'(t) = \mathbf{D}c(t) = \begin{bmatrix} \frac{d\,c_1}{dt}(t) \\ \vdots \\ \frac{d\,c_n}{dt}(t) \end{bmatrix}$$

  - Gradient: for a scalar function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\nabla f(\mathbf{x}) = \mathbf{D}f(\mathbf{x})^{\mathrm{T}} = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}) \end{bmatrix}$$

  - Properties of gradient: $\nabla f(\mathbf{x}_0)$ orthogonal to level sets of f; $\nabla f(\mathbf{x}_0)$ points in direction of fastest increase of f at $\mathbf{x}_0$.
- Properties of the derivative:

- Sum rule:
$$\mathbf{D}(f+g)(\mathbf{x}) = \mathbf{D}f(\mathbf{x}) + \mathbf{D}g(\mathbf{x})$$
- Product rule:
$$\mathbf{D}(fg)(\mathbf{x}) = f(\mathbf{x})\mathbf{D}g(\mathbf{x}) + g(\mathbf{x})\mathbf{D}f(\mathbf{x})$$

- Chain rule:
$$\mathbf{D}(g \circ f)(\mathbf{x}) = \mathbf{D}g(f(\mathbf{x}))\mathbf{D}f(\mathbf{x})$$
(the right hand side represents composition of linear functions; i.e. matrix multiplication)

- Hessian of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$
$$\mathbf{D}^2 f(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}_0) \end{bmatrix}$$

- Taylor approximation to 2nd order:
$$f(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot \left[\mathbf{D}^2 f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\right] + R_2(\mathbf{x}, \mathbf{x}_0)$$
where
$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\|R_2(\mathbf{x}, \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|^2} = 0$$
Note that by letting $\mathbf{x} = \mathbf{x}_0 + \mathbf{h}$, where $\mathbf{h} = (h_1, \ldots, h_n)$, and expanding out, we get the formula from the book:
$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) h_i + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)\, h_i\, h_j + R_2$$

- Extrema (maxima/minima) of functions:
  - Gradient test for local extrema:
  $$f(\mathbf{x}_0) \text{ is an extremum} \implies \nabla f(\mathbf{x}_0) = 0$$
  (But not the converse; i.e. $\nabla f(\mathbf{x}_0) = 0$ does not mean $\mathbf{x}_0$ is necessarily a local max or min)
  - 2nd derivative test for local extrema: if $\nabla f(\mathbf{x}_0) = 0$ and
  $$\mathbf{D}^2 f(\mathbf{x}_0) \text{ positive/negative definite} \implies f(\mathbf{x}_0) \text{ is a local minimum/maximum}$$
  (But not the converse; i.e. if $\mathbf{D}^2 f(\mathbf{x}_0)$ is *not* positive or negative definite, it may still be true that $f(\mathbf{x}_0)$ is a local max or min.)
  - Lagrange multipliers for constrained extrema: in order for f to have an extremum subject to constraints $g_1(\mathbf{x}) = c_1$, ..., $g_k(\mathbf{x}) = c_k$, we must have
  $$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla g_1(\mathbf{x}_0) + \cdots + \lambda_k \nabla g_k(\mathbf{x}_0)$$
  (Combined with the equations $g_1(\mathbf{x}_0) = c_1$, ..., $g_k(\mathbf{x}_0) = c_k$ this gives $n+k$ equations for the $n+k$ unknowns $(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_k)$.)
  - Existence of global maxima and minima for continuous functions
  $$A \subset \mathbb{R}^n \text{ closed, bounded} \implies f : A \to \mathbb{R} \text{ obtains maximum/minimum values}$$
  - Finding global maxima/minima on a closed and bounded set A
    1. Find and list critical points inside A,
    2. Find and list critical points on ∂A,
    3. Take those points for which f obtains largest/smallest values