PROPOSITION C.4
(i) If f and g are convex, then $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ is convex.
(ii) If f is convex, then $h(\mathbf{x}) = af(\mathbf{x})$ is convex for all $a \geq 0$ and is concave for all $a \leq 0$.
(iii) If f is a convex function on a convex set X, then the level set $L(c) = \{\mathbf{x} : \mathbf{x} \in X, f(\mathbf{x}) \leq c\}$ *is a convex set.*
(iv) If $f \in C^1$, then f is convex over a convex set X if and only if $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla \mathbf{f}(\mathbf{x})^\top (\mathbf{y} - \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in X$.
(v) If $f \in C^2$, then f is convex over a convex set X containing an interior point if and only if the Hessian, $\nabla^2 \mathbf{f}(\mathbf{x})$, is positive semidefinite throughout X.

Derivatives and Subderivatives

Let $\mathbf{e}_i$ denote the i^{th} unit vector (the vector with all components zero except for the i^{th} component, which is one). Then the i^{th} *partial derivative* of a function $f : \Re^n \to \Re$ is defined by

$$\frac{\partial}{\partial x_i} f(\mathbf{x}) = \lim_{h \to 0} \frac{1}{h} \left[f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x}) \right],$$

provided the limit exists (where here $h \to 0$ denotes h tending to zero from above or below). If all partial derivatives exist, the *gradient* is defined as the (column) vector

$$\nabla \mathbf{f}(\mathbf{x}) = \left(\frac{\partial}{\partial x_1} f(\mathbf{x}), \ldots, \frac{\partial}{\partial x_n} f(\mathbf{x})\right).$$

If each of the partial derivatives of f at $\mathbf{x}$ is itself a differentiable function of x, then we define the *second partial derivatives* by

$$\frac{\partial}{\partial x_i \partial x_j} f(\mathbf{x}) = \lim_{h \to 0} \frac{1}{h} \left[\frac{\partial}{\partial x_i} f(\mathbf{x} + h\mathbf{e}_j) - \frac{\partial}{\partial x_i} f(\mathbf{x}) \right].$$

The $n \times n$ matrix of second partial derivatives is called the *Hessian* of f at x and is denoted

$$\nabla^2 \mathbf{f}(\mathbf{x}) = \left[\frac{\partial}{\partial x_i \partial x_j} f(\mathbf{x}) \right].$$

Consider a vector direction $\mathbf{d} \in \Re^n$. The *directional derivative* is defined by

$$D_f(\mathbf{x}; \mathbf{d}) = \lim_{h \downarrow 0} \frac{1}{h} \left[f(\mathbf{x} + h\mathbf{d}) - f(\mathbf{x}) \right], \tag{C.1}$$

provide the limit exists. The function f is said to be *differentiable* at $\mathbf{x}$ if and only if $\nabla \mathbf{f}(\mathbf{x})$ exists and

$$D_f(\mathbf{x}; \mathbf{d}) = \nabla \mathbf{f}(\mathbf{x})^\top \mathbf{d}, \quad \forall \mathbf{d} \in \Re^n.$$

A function is said to be *continuously differentiable* on a set X if the gradient $\nabla \mathbf{f}(\mathbf{x})$ exists for all $\mathbf{x} \in X$ and is continuous on X. The class of all continuously differentiable functions on $\Re^n$ is denoted C^1; C^2 denotes the class of all functions with continuous second partial derivatives on $\Re^n$.

For a convex function f, the gradient exists almost everywhere (at all but a countable number of points in X). If f is convex but the gradient does not exist everywhere, it is useful to define a generalization of the gradient called a *subgradient* of f.