Ch.5 Elements of Calculus

5.1 Sequences and Limits

- A number $x^* \in \Re$ is called the limit of the sequence $\{x_k\}$ if for any positive ϵ there is a number such that for all $k > K$, $|x_k - x^*| < \epsilon$

$$x^* = \lim_{k \to \infty} x_k$$

- A sequence that has a limit is called a convergent sequence.
- **Theorem 5.1** *A convergent sequence has only one limit*
- A sequence $\{x^{(k)}\}$ in $\Re^n$ is bounded if there exists a number $B \geq 0$ such that $\|x^{(k)}\| \leq B$ for all $k = 1, 2, \cdots$.
- **Theorem 5.2** *Every convergent sequence is bounded*
- For a sequence $\{x_k\}$ in $\Re$, a number B is called an upper bound (lower bound) if $x_k \leq B$ ($x_k \geq B$) for all $k = 1, 2, \cdots$.
- Any sequence $\{x_k\}$ in $\Re$ that has an upper bound has a least upper bound (also called the supremum), which is the smallest number B that is an upper bound of $\{x_k\}$.
- **Lemma 5.1** *Let $A \in \Re^{n \times n}$. Then, $\lim_{k \to \infty} A^k = O$ if and only if the eigenvalues of A satisfy $|\lambda_i(A)| < 1$, $i = 1, 2, \cdots, n$.*

- **Lemma 5.2** *The series of* $n \times n$ *matrices*

$$I_n + A + A^2 + \cdots + A^k + \cdots$$

converges if and only if $\lim_{k\to\infty} A^k = O$*. In this case the sum of the series equals* $(I_n - A)^{-1}$.

5.2 Differentiability

- A function $A : \Re^n \to \Re^m$ is affine if there exists a linear function $L : \Re^n \to \Re^m$ and a vector $y \in \Re^m$ such that

$$A(x) = L(x) + y$$

for every $x \in \Re^n$.

- A function $f : \Omega \to \Re^m$, $\Omega \subset \Re^n$, is said to be differentiable at $x_o \in \Omega$ if there is an affine function that approximates f near x_o, that is, there exists a linear function $L : \Re^n \to \Re^m$ such that

$$\lim_{x \to x_o, x \in \Omega} \frac{\|f(s) - (L(x - x_0) + f(x_0))\|}{\|x - x_0\|} = 0$$

5.3 The Derivative Matrix

- The matrix L is called the Jacobian matrix, or derivative matrix, of f at x_0, and is denoted $Df(x_0)$.
- Given $f : \Re^n \to \Re$, if ∇f is differentiable, we say that f

is twice differentiable, and the derivative of ∇f as

$$D^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

- The matrix $D^2 f(x)$ is called the Hessian matrix of f at x, and is often also denoted $F(x)$.
- A function $f : \Omega \to \Re^m$, $\Omega \subset \Re^n$, is said to be continuously differentiable on Ω if it is differentiable, and $Df : \Omega \to \Re^{m \times n}$ is continuous. In this case, we write $f \in C^1$.
- If the components of f have continuous partial derivatives of order p, then we write $f \in C^p$.
- Hessian matrix of a function $f : \Re^n \to \Re$ at x is symmetric if f is twice continuously differentiable at x.

5.4 Differentiation Rules

- **Theorem 5.6 (Chain Rule)** *Let $g : D \to \Re$ be differentiable on an open set $D \subset \Re^n$, and let $f : (a, b) \to D$ be differentiable on (a, b). Then, the composite function $h : (a, b) \to \Re$ given by*

$h(t) = g(f(t))$ *is differentiable on* (a,b), *and*

$$h'(t) = Dg(f(t))Df(t) = \nabla g(f(t))^T \begin{bmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{bmatrix}$$

5.5 Level Sets and Gradients

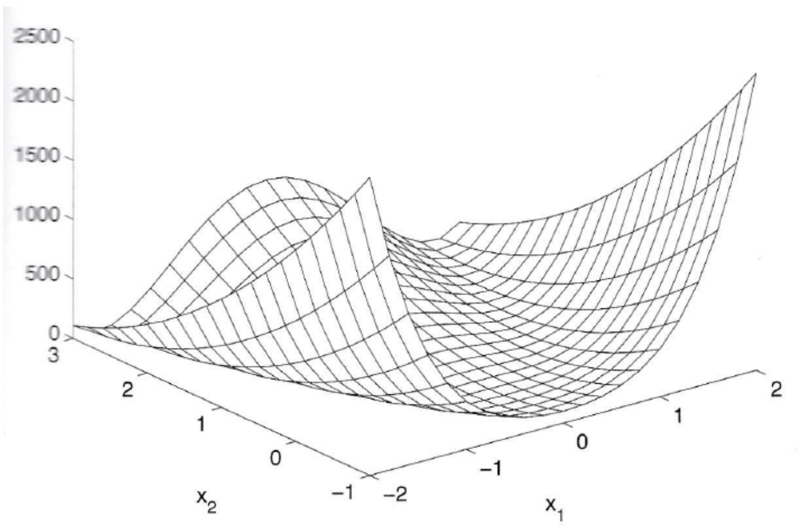
- The level set of a function $f : \Re^n \to \Re$ at level c is the set of points

$$S = \{x : f(x) = c\}$$

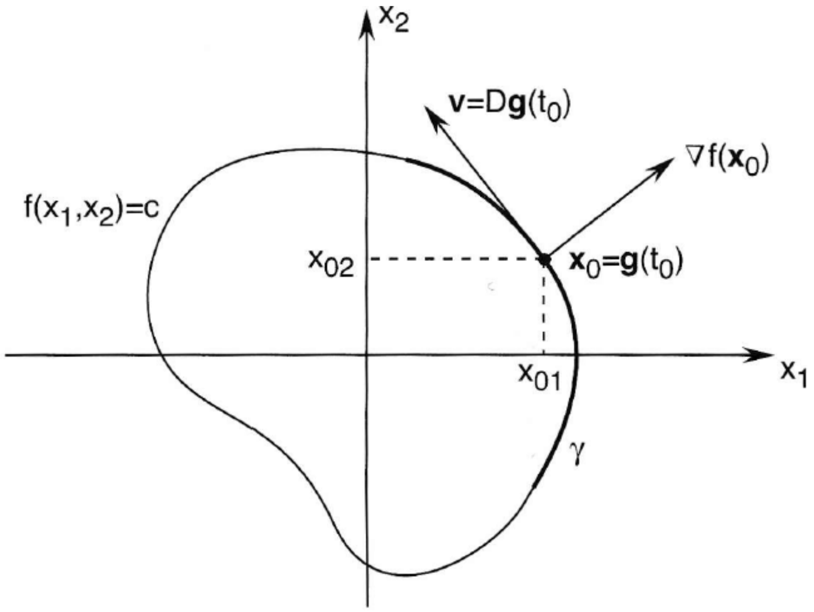
- Example 5.2 consider the following real-valued function on $\Re^2$:

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \;\; x = [x_1, x_2]^T$$

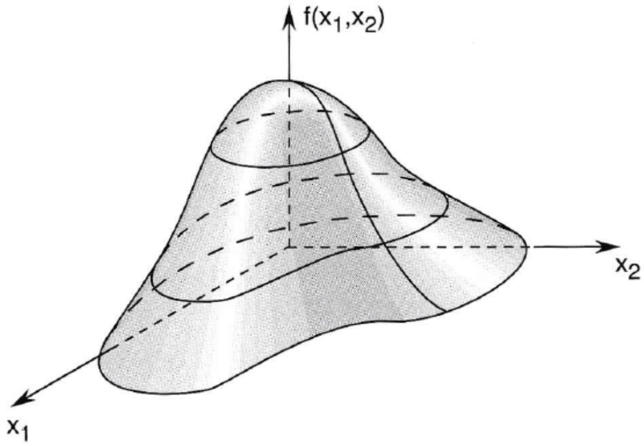
♣ Fig. 5.2 on page 61.

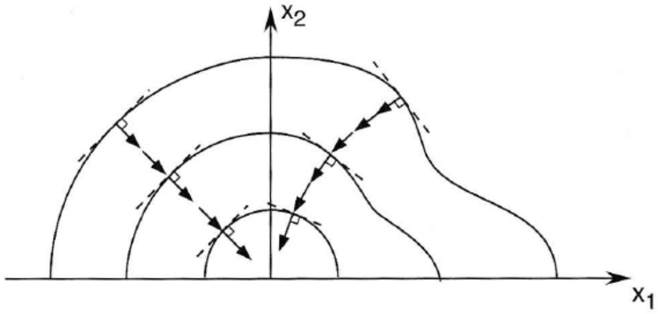


- **Theorem 5.7** *The vector $\nabla f(x_0)$ is orthogonal to the tangent vector to an arbitrary smooth curve passing through x_0 on the level set determined by $f(x) = f(x_0)$.*
  ♣ Fig. 5.4 on page 62.



- $\nabla f(x_0)$ is the direction of maximum rate of increase of f at x_0.
  ♣ Fig. 5.5 on page 63.





5.6 Taylor Series

- **Theorem 5.8 Taylor's theorem** *Assume that a function $f : \Re \to \Re$ is m times continuously differentiable on an interval $[a, b]$. Denote $h = b - a$. Then*

$$\begin{aligned} f(b) &= f(a) + \frac{h}{1!} f^{(1)}(a) + \frac{h^2}{2!} f^{(2)}(a) + \cdots \\ &\quad + \frac{h^{(m-1)}}{(m-1)!} f^{(m-1)}(a) + R_m \end{aligned}$$

where $f^{(i)}$ *is the ith derivative of* f*, and*

$$R_m = \frac{h^m(1-\theta)^{m-1}}{(m-1)!} f^{(m)}(a+\theta h) = \frac{h^m}{m!} f^{(m)}(a+\theta' h)$$

with $\theta, \theta' \in (0,1)$.

- Taylor series expansion of a real-valued function $f : \Re^n \to \Re$ about the point $x_0 \in \Re^n$.

$$\begin{aligned} f(x) &= f(x_0) + \frac{1}{1!} Df(x_0)(x-x_0) \\ &\quad + \frac{1}{2!}(x-x_0)^T D^2 f(x_0)(x-x_0) + o(\|x-x_0\|^2) \end{aligned}$$