Bridging the Gap to Advanced Theorems of Constrained Optimization

Old Theorem

- Lagrange Multiplier Method:
- We want to find the extreme values of a function $f(x,y,z)$ subject to two constraints of the form $g(x,y,z)=k$ and $h(x,y,z)=c$.
- Suppose that the extreme value occurs at (x_0,y_0,z_0) and $\nabla g(x_0,y_0,z_0)$ is not parallel to $\nabla h(x_0,y_0,z_0)$.
- Then at (x_0,y_0,z_0), $\nabla f=\lambda\nabla g+\mu\nabla h$.

Old Theorem

- The method of Lagrange multipliers can be applied to find extreme values of a function of n variables, say $f(\vec{x})$ where $\vec{x}$ is a vector of n variables, $\vec{x}=(x_1,\ldots,x_n)$ subject to $m\le n-1$ constraints ,

$$g_1(\vec{x})=0,\ldots,g_m(\vec{x})=0.$$

Old Theorem

- Assume that f and all of the function g_j have continuous first derivatives in a neighborhood of the point P where the extreme value occurs, and the intersection of the constraint surfaces is smooth near P. Then P is the critical point of the $(n+m)$-variable *Lagrangian function :*

$$L(\vec{x},\lambda_1,\lambda_1,\ldots,\lambda_m)=f(\vec{x})+\sum_{j=1}^{m}\lambda_j g_j(\vec{x})$$

Old Theorem, New Formulation

Theorem 18.2 Question	**Maximize** $f(x_1,x_2,\cdots,x_n)$ **Under constraints** $h(\vec{x})=a_1,\ldots,h_m(\vec{x})=a_m$
Lagrangian Function	$L(x_1,\ldots,x_n,\mu_1,\ldots,\mu_m)=$ $f(\vec{x})-\mu_1[h_1(\vec{x})-a_1]-\cdots-\mu_m[h_m(\vec{x})-a_m]$
Nondegenerate Constraint Qualification (NDCQ)	At the extreme point $\vec{x}^*$, the rank of the $m\times n$ matrix of Jacobian derivatives $Dh(\vec{x}^*)=\left(\frac{\partial h_i}{\partial x_j}\right)_{ij}$ is maximal.
First Order Conditions	There are $\mu_1^*,\ldots,\mu_m^*$ such that for $1\le i\le n$ and $1\le j\le m$ $\frac{\partial L}{\partial x_i}(\vec{x}^*,\mu_1^*,\ldots,\mu_m^*)=0$ $\frac{\partial L}{\partial \mu_j}(\vec{x}^*,\mu_1^*,\ldots,\mu_m^*)=0$

Some New Languages

- Lagrangian Function
- Nondegenerate Constraint Qualification
  - Linear independent vectors
  - Rank of a matrix

Some New Languages

- Definition:
- The vectors $\vec{v}_1, \ldots, \vec{v}_k$ are said to be *linear independent* if the equation

  $$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k = 0$$

  can only be satisfied by $a_1 = a_2 = \cdots = a_k = 0$

Some New Languages

- Definition:
- The vectors $\vec{v}_1, \ldots, \vec{v}_k$ are said to be *linear dependent* if there exists $a_1, a_2, \ldots, a_k$, not all zeros, such that $a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k = 0$.
- Properties:
- If vectors $\vec{v}_1, \ldots, \vec{v}_k$ are linear dependent, then one of the vectors can be written as the linear combination of the others.

Some New Languages

- Definition:
- The *rank* of a $m \times n$ matrix is the dimension of the linear space spanned by the row vectors of the matrix.
- Definition:
- A $m \times n$ matrix ($m < n$) is said to have maximal rank, if the rank is m, which means that the row vectors are linear independent.