## Lecture 4-2

Continuing with the problem we ended with last time, let $A = (a_{ij})$ be an $n \times n$ symmetric matrix, so that $a_{ij} = a_{ji}$ for all indices $i, j$ lying between 1 and $n$. Last time we considered the problem of maximizing or minimizing $f(\vec{x}) = \vec{x}^t A \vec{x}$ for all unit column vectors $x \in \mathbb{R}^n$. Since the set of unit vectors in $\mathbb{R}^n$ is both closed and bounded, we know on general grounds that $f(\vec{x})$ must have both a maximum and a minimum on this set; we saw last time that the maximum and minimum must occur at vectors $\vec{x} \neq \vec{0}$ with $A\vec{x} = \lambda \vec{x}$ for some $\lambda \in \mathbb{R}$; we call any such $\lambda$ an *eigenvalue* of $A$ and the corresponding $\vec{x} \neq \vec{0}$ a $\lambda$-*eigenvector*, or just an *eigenvector*, of $A$. We have thus shown that *$A$ is positive definite in the sense that $\vec{x}^t A \vec{x} > 0$ for all unit vectors $\vec{x}$, and thus for all nonzero vectors $\vec{x}$, if and only if all eigenvalues of $A$ are positive*; similarly we see that $A$ is negative definite in an obvious sense if and only if all of its eigenvalues are negative. We also see that any symmetric $n \times n$ matrix $A$ has a unique largest eigenvalue and a unique smallest eigenvalue (corresponding unit eigenvectors being respective a maximum and a minimum for $f(\vec{x})$ on the (hyper)sphere of radius 1 in $\mathbb{R}^n$, consisting of all unit vectors in that space). Hence if $A$ is positive definite, there are positive constants $c, d$ with $c(\vec{x} \cdot \vec{x}) \leq \vec{x}^t A \vec{x} \leq d(\vec{x} \cdot \vec{x})$ for all $\vec{x} \in \mathbb{R}^n$ (regarding $\vec{x}$ as a column vector). The same result holds with negative constants $c, d$ if instead $A$ is negative definite.

Armed with this new information we return to the situation we discussed on Tuesday of a function $f : \mathbb{R}^n \to \mathbb{R}$ with continuous second partials at a critical point $\vec{a}$. Thanks to the definition of twice differentiability and the second-order Taylor approximation to $f$ near $\vec{a}$, we now see that *if the Hessian matrix $H$ of $f$ at $\vec{a}$ is positive definite, then $f$ has a local minimum at $\vec{a}$, while if $H$ is negative definite at $\vec{a}$, then $f$ has a local maximum at $\vec{a}$.* Also *if $H$ is indefinite in the sense that it has both positive and negative eigenvalues, then $f$ has a saddle point at $\vec{a}$*, since the second-order Taylor approximation shows that $f(\vec{x})$ is less than $f(\vec{a})$ for certain $\vec{x}$ arbitrarily close to $\vec{a}$ (and differing from it by an eigenvector of $H$ with negative eigenvalue), while $f(\vec{x})$ is greater than $f(\vec{a})$ for other $\vec{x}$ arbitrarily close to $\vec{a}$ (and differing from it by an eigenvector of $H$ with positive eigenvalue). If on the other hand one of the eigenvalues of $H$ is 0, then all bets are off, as we cannot say how $f(\vec{x})$ behaves if $\vec{x}$ differs from $\vec{a}$ by an eigenvector of $H$ with 0 eigenvalue.

We look at two more (closely related) problems involving Lagrange multipliers. Consider first the problem of maximizing or minimizing $g(x_1, \ldots, x_n) = a_1 x_1^2 + \ldots + a_n x_n^2$ subject to $x_1^2 + \ldots + x_n^2 = 1$, where the $a_i$ are pairwise distinct constants. Here the Lagrange equation is $(2a_1 x_1, \ldots, 2a_n x_n) = \lambda(2x_1, \ldots, 2x_n)$, which at first seems contradictory, but then we recall that it is possible for some of the $x_i$ to be 0. Returning to the equation with this possibility in mind, we find that the unique maximum of $a_1$ occurs at $(\pm 1, 0 \ldots, 0)$ and the unique minimum of $a_n$ occurs at $(0, \ldots, 0, \pm 1)$, assuming that $a_1$ is the largest of the $a_i$ and $a_n$ is the smallest. Now consider the seemingly easier problem of maximizing $h(\vec{x}) = a_1 x_1 + \ldots, a_n x_n$ under the constraints that $k(\vec{x}) = \sum x_i = 1, x_i \geq 0$ for all $i$ (we add this last constraint so as to guarantee that the set of points under consideration is still closed and bounded). Under the same hypothesis on the $a_i$, we get the same maximum and minimum values as before, namely $a_1$ and $a_n$, respectively, this time at $(1, 0, \ldots, 0)$ and $(0, \ldots, 0, 1)$; but this time the Lagrange equation $(a_1, \ldots, a_n) = \lambda(1, \ldots, 1)$ is definitely *not* satisfied for any choice of $\lambda$. What's going on here? It turns out that this last problem differs in a subtle way from any of the others previously considered: because of the additional constraint that $x_i \geq 0$ for all $i$, the points $(1, 0, \ldots, 0)$ and $(0, \ldots, 0, 1)$ are *not* smooth points of the level set, even though the gradient of the function $k(\vec{x})$ is not $\vec{0}$ at these points. These points are "sharp corners" of the level set and it clearly does not have a well-defined tangent hyperplane at either of them, whence the argument that the Lagrange equation must be satisfied breaks down at these points. The supposedly simpler functions involved in the second problem actually led to a harder situation to analyze. Had we omitted the constraints $x_i \geq 0$, then the Lagrange equation would have correctly told us that $h(\vec{x})$ has no maximum or minimum on any level set of $k$.

Sometimes we have to combine the techniques of §§16.6 and 16.7 in the same problem. Thus if we are asked to maximize $f(x, y) = xy$ on the unit disk $x^2 + y^2 \leq 1$ we first look at the interior of the disk, consisting of all points $(x, y)$ with $x^2 + y^2 < 1$. This is an open set, so the critical points of $f$ on it are the ones with $\nabla f = \vec{0}$. We have already seen that there is just one such point, namely $(0, 0)$, and that it is a saddle point. Next we look at the boundary of the disk, which is exactly the unit circle in the $xy$-plane. Using either Lagrange multipliers or the more elementary method given earlier of parametrizing this boundary, we find that the overall maximum of $f(x, y)$ is $1/2$, attained at $(\sqrt{2}/2, \sqrt{2}/2)$, and the overall minimum of $f(x, y)$ is $-1/2$, attained at $(\sqrt{2}/2, -\sqrt{2}/2)$.