Math 415 - Lecture 36

Minima, maxima and saddle points, Constrained Optimization

Wednesday December 2nd 2015

Textbook reading: Chapter 6.1

Suggested practice exercises: Chapter 6.1, # 1, 16

Strang lecture: Lecture 27: Positive definite matrices and minima

1 Review

Spectral theorem:

- A is a **symmetric** matrix if $A^T = A$.
- Any $n \times n$ symmetric matrix A has n **real eigenvalues** and an **orthonormal eigenbasis** $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$.
- So, we can write $A = QDQ^T$ where

$$D = \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{\text{matrix of eigenvalues}} \quad \text{and } Q = \underbrace{\begin{bmatrix} | & & | \\ \mathbf{q}_1 & \cdots & \mathbf{q}_n \\ | & & | \end{bmatrix}}_{\text{matrix of eigenvectors}}$$

- A is called **positive definite** if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.
- a function of the form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called a **quadratic form**.
- Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A. Then
  1. If all $\lambda_i > 0$, then A is positive definite,
  2. If all $\lambda_i < 0$, then $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$
  3. If some $\lambda_i > 0$, some $\lambda_j < 0$, $\mathbf{x}^T A\mathbf{x}$ will have both positive and negative values.

2 2nd derivative test

2.1 2nd derivative test

Definition 1. Let $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable, the **Hessian** matrix of f

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial^2 x_1}(\mathbf{0}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{0}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{0}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{0}) & \frac{\partial^2 f}{\partial^2 x_2}(\mathbf{0}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{0}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{0}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{0}) & \cdots & \frac{\partial^2 f}{\partial^2 x_n}(\mathbf{0}) \end{bmatrix}$$

Idea. If $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable and $\mathbf{0}$ is a critical point, then $f(\mathbf{x}) \approx f(\mathbf{0}) + \frac{1}{2}\mathbf{x} \cdot H\mathbf{x}$.

- H is always symmetric
- We're approximating $f(\mathbf{x})$ by $f(\mathbf{0})$ plus a quadratic function, $\frac{1}{2}\mathbf{x} \cdot H\mathbf{x}$!
- We understand $q(\mathbf{x}) = \frac{1}{2}\mathbf{x} \cdot H\mathbf{x} \implies$ we understand if $\mathbf{0}$ is a max, min or neither for f!
- Turns out: $q(\mathbf{x})$ is determined by eigenvectors and eigenvalues of H!

Theorem 1 (2nd derivative test)**.** *If $f : \mathbb{R}^n \to \mathbb{R}$ has a critical point at $\mathbf{0}$, then*

1. *If all eigenvalues of H are positive, then $\mathbf{0}$ is a local min.* *H is positive-definite, graph is a bowl.*
2. *If all eigenvalues of H are negative, then $\mathbf{0}$ is a local max.* *H is negative-definite, graph is a dome.*
3. *If one eigenvalue of H is positive and one is negative, then $\mathbf{0}$ is neither a max nor a min.* *H is indefinite, graph is a saddle*
4. *Otherwise (e.g. all eigenvalues positive or zero), no information!*

Example 2. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ has a critical point at $\mathbf{0}$ and has Hessian $H = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Does f have local max, min or neither at $\mathbf{0}$?

(An example of such a function is $f(x, y) = \frac{1}{2}x^2 + 2xy + \frac{1}{2}y^2$).

Solution. We showed that H has eigenvalues 3 and -1. So f has a saddle point at $\mathbf{0}$.

Example 3. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ has a critical point at $\mathbf{0}$ and has Hessian $H = \begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}$. Does f have local max, min or neither at $\mathbf{0}$?

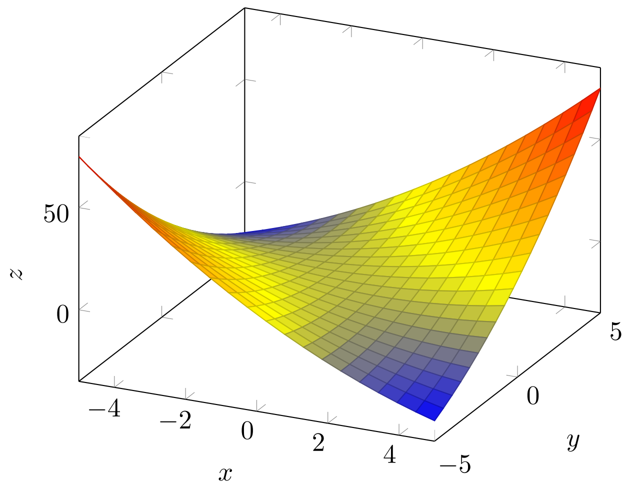

Figure 1: Graph of the function $f(x, y) = \frac{1}{2}x^2 + 2xy + \frac{1}{2}y^2$

Solution. Eigenvalues: Sum $\lambda_1 + \lambda_2 = \text{Tr}(H) = 4$ Product $\lambda_1\lambda_2 = \det(H) = 2$. So, λ_1, λ_2 must be positive! (positive product $\implies$ both positive or both negative. positive sum $\implies$ both positive.)

2nd derivative test says: $f(\mathbf{0})$ is local min.

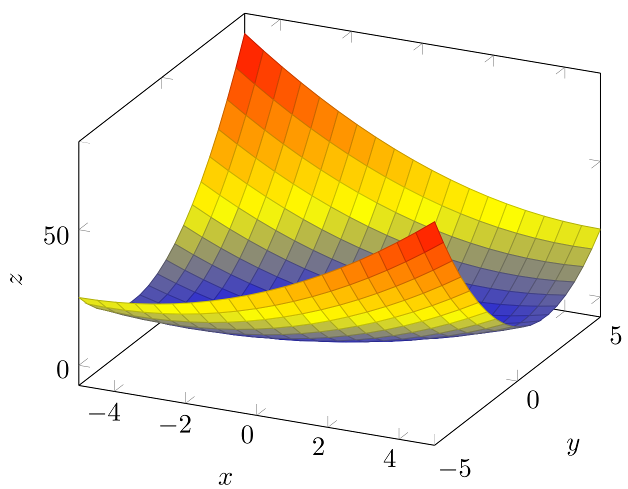

Figure 2: Graph of the function $f(x, y) = \frac{1}{2}x^2 - xy + \frac{3}{2}y^2$

3 Constrained optimization

Problem: Given a quadratic from $q(x)$, find the maximum or minimum value $q(x)$ for $\mathbf{x}$ in some specified set. Typically, the problem can be arranged such that $\mathbf{x}$ varies over the set of vectors with $\mathbf{x}^T\mathbf{x} = 1$.

Example 4. Let $A = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. Find the maximum and minimum values of $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

Solution. The quadratic form is $q(x_1, x_2, x_3) = 9x_1^2 + 4x_2^2 + 3x_3^2$. We are interested in the maximal value for q when (x_1, x_2, x_3) is such that $x_1^2+x_2^2+x_3^2 = 1$. Now we can give an upper bound for q: we obviously have

$$q(\mathbf{x}) \leq 9x_1^2 + 9x_2^2 + 9x_3^2 = 9$$

Solution. So $q(\mathbf{x})$ can not be bigger than 9, for any $\mathbf{x}$. Can we get $q(\mathbf{x}) = 9$ for some $\mathbf{x}$? Obviously for $\mathbf{x} = (1, 0, 0)$ we achieve the upper bound, so 9 is the maximum value for q (under this constraint.) What is a lower bound? For which $\mathbf{x}$ is the lower bound achieved?

What if A is not diagonal?

Theorem 2. *Let A be a symmetric matrix and let λ_m be the least eigenvalue and λ_M be the greatest eigenvalue of A. Then*

$$\lambda_m = \min\{\mathbf{x}^T A\mathbf{x} \ : \ \mathbf{x}^T\mathbf{x} = 1\},$$

moreover if $\mathbf{u}_m$ is a unit eigenvector corresponding to λ_m, then $\mathbf{u}_m^T A\mathbf{u}_m = \lambda_m$. In addition,

$$\lambda_M = \max\{\mathbf{x}^T A\mathbf{x} \ : \ \mathbf{x}^T\mathbf{x} = 1\},$$

moreover if $\mathbf{u}_M$ is a unit eigenvector corresponding to λ_M, then $\mathbf{u}_M^T A\mathbf{u}_M = \lambda_M$.

Proof. We know by the spectral theorem that $A = QDQ^T$, and so we can write $q(\mathbf{x}) = \mathbf{x}^T QDQ^T\mathbf{x} = u^T Du = \lambda_M u_1^2 + \cdots + \lambda_m u_m^2$, where $u = Q^T x$. As before we see that the largest eignvalue λ_M is the upper bound for q, achieved for $u = (1, 0, \ldots, 0)$ or $\mathbf{x} = Qu$, the normalized eigenvector corresponding to λ_M. Same argument for λ_m. □

Example 5. Let $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Find the maximum and minimum values of $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

Solution. We first find eigenvectors and eigenvalues for A. Let us ask Wolfram Alpha: $\det(A)$, Eigenvalues and eigenvectors . So $\lambda = 6, 3, 1$, with eigenvectors

$$v_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Then $q(x)$ has maximum value 6, and $q(v_1) = v_1^T A v_1 = 6\|\mathbf{v_1}\| = 6$. The minimal value is 1 and $q(v_3) = v_3^T A v_3 = 1\|\mathbf{v_3}\| = 1$.