Linear Algebra

Phree

December 13, 2019

Chapter 1

Convex Sets

1.1 Cones

Definition 1.1. A set C is a cone if for every $x \in C$ and $\theta \geq 0$, $\theta x \in C$ always holds.

Cone is also named *nonnegative homogeneous*. It is easy to prove that a cone is closed.

1.2 Dual cones

1.2.1 Dual cones

Definition 1.2. Let K be a cone. The set is called dual cone of K if

$$K^* = \{y | y^{\mathrm{T}} x \geq 0 \text{ for all } x \in K\}.$$

1.2.2 Examples

Dual of a norm cone

The dual cone of $K = \{(x,t) \in \mathbb{R}^{n+1} | \|x\| \leq t\}$ is

$$K^* = \{(u,v) \in \mathbb{R}^{n+1} | \|u\| \leq v\},$$

where the dual norm is given by $\|u\|_* = \sup\{u^T x | \|x\| \leq 1\}$.

Nonnegative orthant

The cone $\mathbb{R}^n_+$ is self-dual.

Positive semi-definitive matrix

The set of positive semi-definitive matrix $\mathbf{S}^n_+$ is also self-dual.

1.2.3 Property of dual cone

- K^* is a cone and convex.
- K is closed and convex.
- $K_1 \subseteq K_2$ implies $K_2^* \subseteq K_1^*$

1.3 Generalized inequalities

1.3.1 Proper cones

Definition 1.3. A cone $K \in \mathbb{R}^n$ is a proper cone if satisfies the following conditions:

- K is convex.
- K is closed.
- K is solid, i.e., it has nonempty interior.
- K is pointed, i.e., it contains no line.

Then we can define generalized inequalities on K, which is a partial ordering on $mathbbR^n$. We write y

Chapter 2

Convex functions

2.1 Convex functions and examples

Definition 2.1. A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if $\mathbf{dom}f$ is a convex set and for all $x, y \in \mathbf{dom}f$, and $\theta \geq 0$, the following equation holds:

$$\theta f(x) + (1-\theta)f(y) \leq f(\theta x + (1-\theta)y) \tag{2.1}$$

It can be shown that a function is convex if and only if it is convex when its domain is restrictecd to a line. A convex function is contiuous on the relative interior of its domain.

2.1.1 First-order condition

Theorem 2.1. *Assume f is differentiable, then f is convex if and only if $\mathbf{dom}f$ is convex and*

$$f(y) \geq f(x) + \nabla f(x)^T(y-x) \tag{2.2}$$

holds for all $x, y \in \mathbf{dom}f$.

2.1.2 Second-order condition

Theorem 2.2. *Supppose that f is twice differentiable and $\mathbf{dom}f$ is open, this is, its* Hession Matrix $\nabla^2 f(x)$ *exists at each point in its domain. Then f is convex if and only if $mathbfdomf$ is convex its Hession Matrix is positive semidefinitite.*

For strict convexity, we have that if $\nabla^2 f(x) \succ 0$ for all $x \in \mathbf{dom}f$, f is strict convex. However, the converse is not true.

2.1.3 Examples

Here are some convex functions on $\mathbb{R}^2$:

- Norms.
- Max function.

- Quadratic-over-linear function, i.e., $f(x, y) = x^2/y$, with
$$\mathbf{dom} f = \mathbb{R} \times \mathbb{R}_{++}.$$
- Log-sum-exp function, i.e., $f(\mathbf{x}) = \log(e^{x_1} + e^{x_2} + \cdots + e^{x_n})$. In fact, this function can be viewed as a differentiable approximation of the max function, since
$$\max(\mathbf{x}) \leq f(\mathbf{x}) \leq \max(\mathbf{x}) + \log(n) \tag{2.3}$$
holds for any $\mathbf{x}$.
- Geometric mean, i.e., $f(x) = (\prod_{i=1}^{n} x_i)^{\frac{1}{n}}$ with $\mathbf{dom} f = \mathbb{R}_{++}$ is concave.
- Log-determinant, i.e., $f(\mathbf{X}) = \log\det \mathbf{X}$ on $\mathbf{dom} f = \mathbb{S}^n_{++}$ is concave. *The proof of this part is sophisticated and I don't fully understand so I should turn back later.*

2.1.4 Sublevel sets

Definition 2.2. The α-sublevel set of a function $f : \mathbb{R}^n \leftarrow \mathbb{R}$ is
$$C_\alpha = \{\mathbf{x} \in \mathbf{dom} f \mid f(\mathbf{x}) < \alpha\}.$$

Sublevel sets of a convex function are convex, while the converse is not true, and this can be used to tell a set is convex by expressing it as a sublevel of a convex function.

2.1.5 Epigraph

Definition 2.3. The epigraph of a function $f : \mathbb{R}^n \leftarrow \mathbb{R}$ is
$$\mathbf{epi} = \{(\mathbf{x}, t)) \mid \mathbf{x} \in \mathbf{dom} f, f(\mathbf{x}) \leq t\}.$$

The relation between convex function and convex set is that *a function is convex if and only if its epigraph is a convex set.*

Chapter 3

Gradient Methods

3.1 Conjugate Gradient Method

Consider minimizing the quadratic form

$$f(x) = \frac{1}{2}x^T Gx + b \tag{3.1}$$

where $G \in \mathbf{S}_{++}^n$ and $x \in \mathbb{R}^n$. First let's restrict the case to $n = 2$. If G is a diagonal matrix, then intuitively we can find the solution of 3.1 in two steps by exact line searching along with two coordinates, since the contour of 3.1 is a ellipse with its major and minor axes parallel to the coordinates.

In general case, we can diagonalize G by introducing a linear transform, denoted by $D = (d_1, \cdots, d_n)$. Substituting $x = D\tilde{x}$ for 3.1 we get

$$\tilde{f}(\tilde{x}) = f(Dx) = \frac{1}{2}x^T D^T GDx. \tag{3.2}$$

The (i, j)th entry of $D^T GD$ is $d_i^T Gd_j$ and $d_i^T Gd_j = 0$ if $i \neq j$. Here we encounter a very interesting condition, which can be denoted as

$$\langle d_i, d_j \rangle_G = d_i^T Gd_j = 0, \tag{3.3}$$

if we introduce a new inner product

$$\langle x, y \rangle_A = x^T Ay. \tag{3.4}$$

where $A \in \mathbf{S}_{++}^n$. Then condition 3.3, called conjugate, can be viewed as a generalization of orthogonality.

Definition 3.1. Let $G \in \mathbf{S}_{++}^n$ and $x, y \in \mathbb{R}^n$ are nonzero vectors. x and y is conjugate with respect to G if

$$\langle x, y \rangle_G = x^T Gy = 0. \tag{3.5}$$

A sequence of vectors $\{d_i\}$, $i = 1, 2, \cdots, m$ is conjugate if any two vectors of them are conjugate.

We have the following simple theorem, of which the proof is straightforward.

Theorem 3.1. *Conjugate vectors are linear independent.*

3.1.1 Connection with linear equations