Higher-Order Newton Method for Mathematical Optimization

by

Lucas ter Voert (SNR: 2081136)

A thesis submitted in partial fulfillment of the requirements for the degree of Master in Business Analytics and Operations Research

Tilburg School of Economics and Management
Tilburg University

Supervised by: Prof. Dr. Etienne de Klerk

Date: July 4, 2025

Contents

Acknowledgement

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Etienne de Klerk, for guiding me through the process of writing my master's thesis and for offering me a PhD position in the field of optimization. Thank you for your extremely helpful insights, valuable feedback, and prompt responses. I look forward to collaborating with you on further research in the coming years at Tilburg University.

Secondly, I would like to thank my parents, Rico and Wendy, for their financial and moral support throughout my studies, and for their patience during the many long days I spent studying in my room. Completing my master's program within one year required a great deal of energy and dedication, but I promise to be more present and social at home very soon.

I am also grateful to all my friends and fellow members of Asset | Econometrics. You all made my time at Tilburg University more enjoyable, and I promise to stay in close contact with you during my years as a PhD student.

Lastly, I would like to dedicate a few words to my dear girlfriend, Renate, who obtained her master's degree in Operations Research just last week at the time of writing. We met only a few months ago, while we were both working on our master's theses, so we sometimes had different priorities than seeing each other, but we always understood. You mean a lot to me.

Chapter 1

Introduction

In many real-world situations, an agent seeks to minimize a cost or maximize a reward by taking an optimal action subject to a set of constraints. Such a problem is called an *optimization problem*. In this thesis, we analyze a particular class of optimization problems known as *mathematical* optimization problems. In a mathematical optimization problem, we minimize a real-valued objective function over a feasible region defined by finitely many constraints.

For unconstrained optimization problems, the two most widely used solution algorithms are *gradient descent* and *Newton's method*. Both algorithms start from an initial point and aim to construct a sequence that converges to a minimizer of the objective function.

- Gradient descent updates the current iterate by moving in the direction opposite to the gradient of the objective function, aiming to decrease its value.

- Newton's method iteratively minimizes second-order (quadratic) Taylor approximations of the objective function, aiming to move closer to a minimizer. Note that a quadratic function has a unique minimizer with a closed-form solution if its Hessian is positive definite.

Under certain assumptions, gradient descent or Newton's method converges to a (local) minimizer of the objective function. For example, if the initial point is sufficiently close to a local minimizer, then Newton's method converges quadratically fast to that point.

In this thesis, we study a higher-order Newton method that achieves improved convergence rates compared to classical gradient descent and Newton's method. We then extend this algorithm to solve optimization problems with constraints.

1.1 Previous research

Nesterov and Polyak (2006) developed a variation of Newton's method that incorporates a cubic regularization term into the second-order Taylor approximation of the objective function at each iteration. They proved that, similarly to the classical Newton method, the method with cubic regularization exhibits local quadratic convergence. Moreover, despite the fact that a minimizer of the subproblem's objective function need not be unique and may not have a closed-form solution, Nesterov and Polyak (2006) showed that a global minimizer can be found using linear algebraic techniques.

Lasserre (2001) introduced the *Lasserre hierarchy*, a sequence of semidefinite programs (SDPs) designed to solve *polynomial* optimization problems, i.e., optimization problems with polynomial objective functions and polynomial constraints. This work was later extended by Lasserre (2009) and De Klerk and Laurent (2011). A key result from these papers is that the Lasserre hierarchy converges in a single step, i.e., the problem can be solved by solving a single SDP, if the objective function and constraints are *SOS-convex*.

Ahmadi et al. (2024) developed a variation of Newton's method that minimizes Taylor approximations of order higher than two. Inspired by Nesterov and Polyak (2006), they added a regularization term such that the regularized Taylor approximation becomes an SOS-convex polynomial. They then applied the Lasserre hierarchy to solve the subproblem. Ahmadi et al. (2024) proved that their higher-order Newton method converges locally at a rate faster than quadratic.

Similar to the higher-order Newton method, Doikov and Nesterov (2021) and Doikov and Nesterov (2024) used methods based on *tensors* of higher-order derivatives to minimize a *composite* objective function, i.e., a function that can be written as the sum of a smooth and a nonsmooth term. Under convexity and Lipschitz assumptions, their tensor methods achieve local convergence at a *superlinear* rate, i.e., faster than linear convergence. It is worth noting that these papers predate Ahmadi et al. (2024). A recent paper by Zhu and Cartis (2024) discusses the SOS-convex regularization techniques from Ahmadi et al. (2024) and analyzes global convergence for general nonconvex objective functions.

1.2 Organization and contribution

In Chapter 2, we formally define a mathematical optimization problem and explain the mathematics underlying the gradient descent and Newton's method algorithms.

In Chapter 3, we introduce multivariate polynomials and Taylor approximations. We then present the notions of *SOS* and *SOS-convex* polynomials, and prove some important properties in Chapter 4. In Chapter 5, we present the theory of polynomial optimization from Lasserre (2009) and De Klerk and Laurent (2011) necessary for our purposes.

In Chapter 6 we discuss theory on algebraically open sets and conic decompositions of vector spaces, which is essential for proving important results in Chapter 7, where we present the higher-order Newton method of Ahmadi et al. (2024) and prove its convergence properties.

Our main contribution is the extension of the higher-order Newton method to optimization problems with SOS-convex polynomial constraints, which we present in Section 7.6.

Chapter 2

Mathematical optimization

2.1 Mathematical optimization problems

We define a mathematical optimization problem as follows.

Definition 2.1 (Mathematical optimization problem)**.** A *mathematical optimization problem* is an optimization problem

$$\begin{array}{ll} \inf\limits_{x\in D} & f\left(x\right) \\ \text{s.t.} & g_i\left(x\right) \leq 0, \quad i \in \{1,\ldots,m\} \end{array} \tag{P}$$

where $f, g_1, \ldots, g_m$ are functions on an open domain $D \subseteq \mathbb{R}^n$.

The function f is called the *objective function* of (P). The conditions $g_1\left(x\right) \leq 0, \ldots, g_m\left(x\right) \leq 0$ are called the *constraints*. The set of points satisfying all constraints,

$$T = \{x \in D : g_1\left(x\right) \leq 0, \ldots, g_m\left(x\right) \leq 0\},$$

is called the *feasible set* or *feasible region* of (P).

The problem is to find a point $x^* \in T$ such that

$$f\left(x^*\right) = \inf\left\{f\left(x\right) : x \in T\right\}.$$

A class of mathematical optimization problems that are relatively easy to solve is the class of so-called *convex* optimization problems. A convex optimization problem is a problem with a convex objective function and convex constraints.

Definition 2.2 (Convex optimization problem)**.** A *convex optimization problem* is an optimization problem

$$\begin{aligned} \inf_{x\in D} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \le 0, \quad i \in \{1,\dots,m\} \end{aligned} \tag{CVX}$$

where $f, g_1, \dots, g_m : D \to \mathbb{R}$ are convex functions on an open convex domain $D \subseteq \mathbb{R}^n$.

It follows directly that the feasible region of a convex optimization problem is a convex set.

Theorem 2.1. Let T be the feasible region of (CVX), i.e.,

$$T = \{x \in D : g_1(x) \le 0, \dots, g_m(x) \le 0\}.$$

Then, T is a convex set.

Proof. Let $x, y \in T$ and $\lambda \in [0,1]$. Since D is a convex set, we then have $\lambda x + (1-\lambda) y \in D$. Moreover, since $g_1, \dots, g_m$ are convex functions, it follows that, for all $i \in \{1,\dots,m\}$, we have

$$g_i(\lambda x + (1-\lambda) y) \le \lambda g_i(x) + (1-\lambda) g_i(y) \le \lambda \cdot 0 + (1-\lambda) \cdot 0 = 0.$$

Therefore, $\lambda x + (1-\lambda) y \in T$. Hence, T is a convex set. □

The main reason why a convex optimization problem is easier to solve than a general mathematical optimization problem is that any local minimizer of a convex optimization problem is automatically a global minimizer.

Theorem 2.2. Assume that (CVX) has a local minimizer x^*. Then, x^* is also a global minimizer of (CVX).

Proof. Suppose that x^* is not a global minimizer of (CVX). Let T be the feasible region of (CVX). Then, there exists a $y^* \in T$ such that $f(y^*) < f(x^*)$. Since T is a convex set, and f a convex function, it follows that, for all $\lambda \in [0,1)$, we have $\lambda x^* + (1-\lambda) y^* \in T$, and

$$f(\lambda x^* + (1-\lambda) y^*) \le \lambda f(x^*) + (1-\lambda) f(y^*) < f(x^*).$$

However, x^* would then not be a local minimizer of (CVX). Hence, if x^* is a local minimizer of (CVX), then x^* is also a global minimizer of (CVX). □

2.2 Gradient descent and Newton's method

Consider an unconstrained optimization problem

$$\inf_{x \in D} f(x),$$

where $f : D \to \mathbb{R}$ is a function on an open domain $D \subseteq \mathbb{R}^n$.

Gradient descent and *Newton's method* are solution algorithms that start from an initial point $x_0 \in D$ and aim to construct a sequence $(x_k)_{k=0}^{\infty}$ that converges to a (local) minimizer x^* of f.

Assuming that f is differentiable, gradient descent updates the current iterate x_k by moving in the direction opposite to the gradient ∇f. That is, for some *step size* $t > 0$, we set

$$x_{k+1} = x_k - t \nabla f(x_k).$$

We assume that $x_{k+1} \in D$. Note that, if $\nabla f(x_k) \neq 0$, then there exists a $\delta > 0$ such that, if $t < \delta$, then $f(x_{k+1}) < f(x_k)$.

Since $\nabla f(x^*) = 0$, we terminate the algorithm when $\|\nabla f(x_k)\| < \varepsilon$ for some small *tolerance parameter* $\varepsilon > 0$. When this condition is satisfied, we say that the algorithm has *converged*.

Algorithm 2.1: Gradient descent

Input: Objective function f, initial point x_0, step size $t > 0$, tolerance parameter $\varepsilon > 0$, maximum number of iterations N

Output: Approximate minimizer of f

```
for k ← 0 to N − 1 do
    if ‖∇f(x_k)‖ < ε then
        return x_k // Converged
    x_{k+1} ← x_k − t∇f(x_k)
return x_N // Maximum number of iterations reached
```

Assuming that f is twice continuously differentiable, Newton's method updates the current iterate x_k by minimizing the second-order Taylor approximation of f at x_k, which is given by the quadratic function

$$x \mapsto f(x_k) + (\nabla f(x_k))^\top (x - x_k) + \frac{1}{2}(x - x_k)^\top \nabla^2 f(x_k)(x - x_k).$$

We assume that $\nabla^2 f(x_k) \succ 0$. Then, this function is minimized by setting

$$x_{k+1} = x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k).$$

Algorithm 2.2: Newton's method

Input: Objective function f, initial point x_0, tolerance parameter $\varepsilon > 0$, maximum number of iterations N

Output: Approximate minimizer of f

```
for k ← 0 to N − 1 do
    if ‖∇f(x_k)‖ < ε then
        return x_k // Converged
    x_{k+1} ← x_k − (∇²f(x_k))^{-1} ∇f(x_k)
return x_N // Maximum number of iterations reached
```

We prove the convergence results of gradient descent and Newton's method in Appendix B.

Chapter 3

Multivariate polynomials and Taylor approximations

3.1 Multi-index notation

Multivariate polynomials of degree at most two, and partial derivatives of order at most two, are typically expressed in matrix notation. However, for higher degrees or orders, this approach becomes cumbersome. Therefore, we instead use *multi-index* notation.

We denote the set of nonnegative integers by $\mathbb{N}_0$. That is, $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. We can now define a multi-index.

Definition 3.1 (Multi-index)**.** A *multi-index* is a tuple

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n.$$

We denote the sum of the elements of a multi-index $\alpha \in \mathbb{N}_0^n$ by $|\alpha|$.

Definition 3.2. Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$. Then,

$$|\alpha| = \alpha_1 + \cdots + \alpha_n.$$

We denote the set of multi-indices $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq d$ by $\mathbb{N}_d^n$.

Definition 3.3. $\mathbb{N}_d^n = \{\alpha \in \mathbb{N}_0^n : |\alpha| \leq d\}$.

The number of elements of $\mathbb{N}_d^n$ is given by the following theorem.

Theorem 3.1. $|\mathbb{N}_d^n| = \binom{n+d}{d}$.

Proof. We have

$$|\mathbb{N}_d^n| = |\{\alpha \in \mathbb{N}_0^n : |\alpha| \le d\}| = \left|\left\{(\alpha, k) \in \mathbb{N}_0^{n+1} : |\alpha| + k = d\right\}\right|.$$

Therefore, $|\mathbb{N}_d^n|$ is the number of combinations with repetition of d elements from a set of $n+1$ elements. Hence,

$$|\mathbb{N}_d^n| = \binom{(n+1)+d-1}{d} = \binom{n+d}{d}.$$

□

We introduce the following notation.

Definition 3.4 (Factorial)**.** Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$. Then,

$$\alpha! = \alpha_1! \cdots \alpha_n!.$$

Definition 3.5 (Multinomial coefficient)**.** Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ with $|\alpha| = \alpha_1 + \cdots + \alpha_n = d$. Then,

$$\binom{d}{\alpha} = \binom{d}{\alpha_1, \dots, \alpha_n}.$$

Definition 3.6 (Multi-binomial coefficient)**.** Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ and $\nu = (\nu_1, \dots, \nu_n) \in \mathbb{N}_0^n$ with $\nu \le \alpha$, or $\nu_1 \le \alpha_1, \dots, \nu_n \le \alpha_n$. Then,

$$\binom{\alpha}{\nu} = \binom{\alpha_1}{\nu_1} \cdots \binom{\alpha_n}{\nu_n}.$$

Definition 3.7 (Monomial)**.** Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$. Then,

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

We can now express the multinomial theorem in multi-index notation.

Theorem 3.2 (Multinomial theorem; e.g., Folland, 2002, Theorem 2.52)**.** For all $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, we have

$$(x_1 + \cdots + x_n)^d = \sum_{\substack{\alpha \in \mathbb{N}_0 \\ |\alpha| = d}} \binom{d}{\alpha} x^\alpha.$$

We can also generalize the binomial theorem to multivariate cases.

Theorem 3.3 (Multi-binomial theorem; e.g., Saint Raymond, 1991, Theorem 1.2)**.** For all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}_0^n$, we have

$$(x+y)^\alpha = \sum_{\substack{\nu \in \mathbb{N}_0^n \\ \nu \leq \alpha}} \binom{\alpha}{\nu} x^\nu y^{\alpha-\nu}.$$

Proof. Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, and $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$. From the binomial theorem, it then follows that

$$\begin{aligned}
(x+y)^\alpha &= (x_1+y_1)^{\alpha_1} \cdots (x_n+y_n)^{\alpha_n} \\
&= \sum_{\nu_1=0}^{\alpha_1} \binom{\alpha_1}{\nu_1} x_1^{\nu_1} y_1^{\alpha_1-\nu_1} \cdots \sum_{\nu_n=0}^{\alpha_n} \binom{\alpha_n}{\nu_n} x_n^{\nu_n} y_n^{\alpha_n-\nu_n} \\
&= \sum_{\nu_1=0}^{\alpha_1} \cdots \sum_{\nu_n=0}^{\alpha_n} \binom{\alpha_1}{\nu_1} \cdots \binom{\alpha_n}{\nu_n} x_1^{\nu_1} \cdots x_n^{\nu_n} y_1^{\alpha_1-\nu_1} \cdots y_n^{\alpha_n-\nu_n} \\
&= \sum_{\substack{\nu \in \mathbb{N}_0^n \\ \nu \leq \alpha}} \binom{\alpha}{\nu} x^\nu y^{\alpha-\nu}.
\end{aligned}$$

□

3.2 Multivariate polynomials

We can now express multivariate polynomials in multi-index notation.

Definition 3.8 (Polynomial)**.** An *n-variate polynomial of degree d* is a function $p : \mathbb{R}^n \to \mathbb{R}$, defined for all $x \in \mathbb{R}^n$ by

$$p(x) = \sum_{\alpha \in \mathbb{N}_d^n} p_\alpha x^\alpha,$$

where $(p_\alpha)_{\alpha \in \mathbb{N}_d^n} \in \mathbb{R}^{\binom{n+d}{d}}$.

It follows directly from this definition that the set of n-variate polynomial of degree d is a vector space.

Theorem 3.4. The set of n-variate polynomials of degree d is an $\binom{n+d}{d}$-dimensional vector space with basis

$$\{x \mapsto x^\alpha : \alpha \in \mathbb{N}_d^n\}.$$

We denote the vector of basic monomials of degree at most d by $x \mapsto [x]_d$.

Definition 3.9. Let $x \in \mathbb{R}^n$. Then,

$$[x]_d = (x^\alpha)_{\alpha \in \mathbb{N}^n_d}.$$

It follows directly that an n-variate polynomial of degree d can be uniquely represented by a vector $c \in \mathbb{R}^{\binom{n+d}{d}}$.

Corollary 3.4.1. Let p be an n-variate polynomial of degree d. Then, there exists a unique $c \in \mathbb{R}^{\binom{n+d}{d}}$ such that, for all $x \in \mathbb{R}^n$, we have

$$p(x) = c^\top [x]_d.$$

If d is even, then an n-variate polynomial of degree d can also be represented by a matrix $Q \in \mathbb{S}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}}}$. Note that, in general, Q is not unique.

Theorem 3.5. Let p be an n-variate polynomial of an even degree d. Then, there exists a $Q \in \mathbb{S}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}}}$ such that, for all $x \in \mathbb{R}^n$, we have

$$p(x) = [x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}}.$$

Proof. Let $Q = (q_{\beta,\gamma})_{\beta,\gamma \in \mathbb{N}^n_{\frac{d}{2}}} \in \mathbb{S}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}}}$. For all $x \in \mathbb{R}^n$, we then have

$$[x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}} = \sum_{\beta \in \mathbb{N}^n_{\frac{d}{2}}} \sum_{\gamma \in \mathbb{N}^n_{\frac{d}{2}}} q_{\beta,\gamma} x^{\beta+\gamma} = \sum_{\alpha \in \mathbb{N}^n_d} \left(\sum_{\substack{\beta \in \mathbb{N}^n_{\frac{d}{2}} \\ \beta+\gamma=\alpha}} \sum_{\gamma \in \mathbb{N}^n_{\frac{d}{2}}} q_{\beta,\gamma} \right) x^\alpha.$$

Hence, if $\sum_{\substack{\beta \in \mathbb{N}^n_{\frac{d}{2}} \\ \beta+\gamma=\alpha}} \sum_{\gamma \in \mathbb{N}^n_{\frac{d}{2}}} q_{\beta,\gamma} = p_\alpha$ for all $\alpha \in \mathbb{N}^n_d$, then $p(x) = [x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}}$ for all $x \in \mathbb{R}^n$. $\square$

In this thesis, we use the following inner product and norm on a vector space of polynomials.

Definition 3.10 (Polynomial inner product and norm)**.** We define the inner product on the vector space V of n-variate polynomials of degree d for all $p, q \in V$ by

$$\langle p, q \rangle = \sum_{\alpha \in \mathbb{N}_d^n} p_\alpha q_\alpha,$$

and the norm on V for all $p \in V$ by

$$\|p\| = \sqrt{\langle p, p \rangle} = \sqrt{\sum_{\alpha \in \mathbb{N}_d^n} p_\alpha^2}.$$

3.3 Partial derivatives

We now express partial derivatives in multi-index notation.

Definition 3.11 (Partial derivative)**.** Let $f : D \to \mathbb{R}^n$ be a d times partially differentiable function on an open domain $D \subseteq \mathbb{R}^n$. Then, the *partial derivative* $D^\alpha f$ of f with respect to a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_d^n$ is defined for all $x = (x_1, \ldots, x_n) \in D$ by

$$D^\alpha f(x) = \frac{\partial^{|\alpha|} f(x)}{\partial x^\alpha} = \frac{\partial^{\alpha_1 + \cdots + \alpha_n} f(x_1, \ldots, x_n)}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

For differentiating univariate monomials, we can use the following theorem.

Theorem 3.6 (Power rule)**.** For all $x \in \mathbb{R}$ and $k, l \in \mathbb{N}_0$, we have

$$\frac{\partial^k x^l}{\partial x^k} = \begin{cases} \frac{l!}{(l-k)!} x^{l-k} & \text{if } k \leq l \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let $x \in \mathbb{R}^n$ and $l \in \mathbb{N}_0$. Then, $\frac{\partial^0 x^l}{\partial x^0} = x^l$. If, for some $k \in \mathbb{N}_0$, we have

$$\frac{\partial^k x^l}{\partial x^k} = \begin{cases} \frac{l!}{(l-k)!} x^{l-k} & \text{if } k \leq l \\ 0 & \text{otherwise,} \end{cases}$$

then

$$\frac{\partial^{k+1} x^l}{\partial x^{k+1}} = \frac{\partial}{\partial x} \left(\frac{\partial^k x^l}{\partial x^k} \right) = \begin{cases} \frac{l!}{(k-(k+1))!} x^{l-(k+1)} & \text{if } k+1 \leq l \\ 0 & \text{otherwise.} \end{cases}$$

Hence, for all $k \in \mathbb{N}_0$, we have

$$\frac{\partial^k x^l}{\partial x^k} = \begin{cases} \frac{l!}{(l-k)!} x^{l-k} & \text{if } k \leq l \\ 0 & \text{otherwise.} \end{cases}$$

□

We can generalize the power rule to multivariate cases.

Theorem 3.7 (Multivariate power rule)**.** For all $x \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{N}_0^n$, we have

$$\frac{\partial^{|\alpha|} x^\beta}{\partial x^\alpha} = \begin{cases} \frac{\beta!}{(\beta-\alpha)!} x^{\beta-\alpha} & \text{if } \alpha \leq \beta \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$, and $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}_0^n$. From the univariate power rule, it then follows that, for all $i \in \{1, \dots, n\}$, we have

$$\frac{\partial^{\alpha_i} x_i^{\beta_i}}{\partial x_i^{\alpha_i}} = \begin{cases} \frac{\beta_i!}{(\beta_i-\alpha_i)!} x_i^{\beta_i-\alpha_i} & \text{if } \alpha_i \leq \beta_i \\ 0 & \text{otherwise.} \end{cases}$$

Hence,

$$\frac{\partial^{|\alpha|} x^\beta}{\partial x^\alpha} = \frac{\partial^{\alpha_1+\dots+\alpha_n} x_1^{\beta_1} \cdots x_n^{\beta_n}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} = \frac{\partial^{\alpha_1} x_1^{\beta_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n} x_n^{\beta_n}}{\partial x_n^{\alpha_n}} = \begin{cases} \frac{\beta!}{(\beta-\alpha)!} x^{\beta-\alpha} & \text{if } \alpha \leq \beta \\ 0 & \text{otherwise.} \end{cases}$$

□

3.4 Taylor approximations

A d times continuously differentiable function $f : D \to \mathbb{R}$ on an open domain $D \subseteq \mathbb{R}^n$ can be locally approximated around a point $x \in D$ by a polynomial of degree d, called the d^{th}*-order Taylor approximation* of f at x.

Definition 3.12 (Taylor approximation)**.** Let $f : D \to \mathbb{R}$ be a d times continuously differentiable function on an open domain $D \subseteq \mathbb{R}^n$. Then, the d^{th}*-order Taylor approximation* $T_{x,d} : \mathbb{R}^n \to \mathbb{R}$ of f, at a point $x \in D$, is defined for all $y \in \mathbb{R}^n$ by

$$T_{x,d}(y) = \sum_{\alpha \in \mathbb{N}_d^n} \frac{D^\alpha f(x)}{\alpha!} (y - x)^\alpha .$$

The remainder $R_{x,d} : \mathbb{R}^n \to \mathbb{R}$ of $T_{x,d}$ is defined for all $y \in D$ by

$$R_{x,d}(y) = f(y) - T_{x,d}(y) .$$

If the d^{th}-order partial derivatives of f are Lipschitz continuous (see Definition B.1), then we can find bounds on $R_{x,d}$.

Theorem 3.8 (e.g., Folland, 2002, Section 2.7)**.** Let $f : D \to \mathbb{R}^n$ be a d times continuously differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that, for all $\alpha \in \mathbb{N}_d^n$ with $|\alpha| = d$, the function $D^\alpha f$ is Lipschitz continuous with Lipschitz constant L. Let $k \in \{0, \dots, d\}$. For all $\alpha \in \mathbb{N}_d^n$ with $|\alpha| = d - k$, and $x, y \in D$, we then have

$$|D^\alpha R_{x,d}(y)| \leq \frac{\sqrt{n^k} L}{(k+1)!} \|y - x\|^{k+1}.$$

Proof. We prove the theorem by induction on k. Let $\alpha \in \mathbb{N}_d^n$ with $|\alpha| = d$, and $x, y \in D$. From the power rule, it then follows that

$$D^\alpha T_{x,d}(y) = \sum_{\substack{\beta \in \mathbb{N}_d^n \\ \beta \geq \alpha}} \frac{D^\beta f(x)}{(\beta - \alpha)!} (y - x)^{\beta - \alpha} = D^\alpha f(x).$$

From the Lipschitz property, it now follows that

$$|D^\alpha R(y)| = |D^\alpha f(y) - D^\alpha T(y)| = |D^\alpha f(y) - D^\alpha f(x)| \leq L \|y - x\|,$$

so the theorem is true for $k = 0$.

Now, assume that the theorem is true for some $k \in \{0, \dots, d-1\}$. Let $\alpha \in \mathbb{N}_d^n$ with $|\alpha| = d - k$. From the power rule, it then follows that, for all $x, y \in D$, we have $D^\alpha T_{x,d}(y) = \sum_{\substack{\beta \in \mathbb{N}_d^n \\ \beta \geq \alpha}} \frac{D^\alpha f(x)}{(\beta - \alpha)!} (y - x)^{\beta - \alpha}$. Therefore, for all $x \in D$, we have $D^\alpha T_{x,d}(x) = D^\alpha f(x)$, so $D^\alpha R_{x,d}(x) = 0$.

Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. From the chain rule and the fundamental theorem of calculus, it then follows that

$$\begin{aligned}
|D^\alpha R_{x,d}(y)| &= \left| \int_0^1 \sum_{i=1}^n D^{\alpha + e_i} R_{x,d}(x + t(y - x))(y_i - x_i)\, dt \right| \\
&\leq \int_0^1 \sum_{i=1}^n \left| D^{\alpha + e_i} R_{x,d}(x + t(y - x)) \right| |y_i - x_i|\, dt \\
&\leq \int_0^1 \sum_{i=1}^n \frac{t^{k+1} \sqrt{n^k} L}{(k+1)!} \|y - x\|^{k+1} |y_i - x_i|\, dt \\
&\leq \int_0^1 \frac{t^{k+1} \sqrt{n^{k+1}} L}{(k+1)!} \|y - x\|^{k+2}\, dt = \frac{\sqrt{n^{k+1}} L}{(k+2)!} \|y - x\|^{k+2},
\end{aligned}$$

where the first inequality follows from the triangle inequality, the second from our induction hypothesis, and third from the Cauchy-Schwarz inequality. Therefore, the theorem is also true for $k + 1$.

Hence, the theorem is true for all $k \in \{0, \dots, d\}$. □

Chapter 4

SOS and SOS-convex polynomials

4.1 SOS polynomials

In general, determining whether an a polynomial is nonnegative is NP-hard (e.g., Ahmadi et al., 2024). However, a sufficient condition is that it is a *sum of squares*.

Definition 4.1 (SOS polynomial)**.** A polynomial p of an even degree d is called a *sum of squares (SOS)* if there exist polynomials $q_1, \ldots, q_r$ of degree $\frac{d}{2}$ such that

$$p = \sum_{i=1}^{r} q_i^2.$$

Note that every SOS polynomial is nonnegative. However, not every nonnegative polynomial is also SOS (e.g., Ahmadi et al., 2024).

Recall from Theorem 3.5 that a polynomial p of an even degree can be represented by a matrix. Now, p is SOS if and only if it can be represented by a positive-semidefinite matrix.

Theorem 4.1 (e.g., Ahmadi et al., 2024)**.** An n-variate polynomial p of an even degree d is SOS if and only if there exists a $Q \succeq 0$ such that, for all $x \in \mathbb{R}^n$, we have

$$p(x) = [x]_{\frac{d}{2}}^{\top} Q \, [x]_{\frac{d}{2}}.$$

Proof. $\Longrightarrow$) Assume that p is SOS. Then, there exist n-variate polynomials $q_1, \ldots, q_r$ of degree $\frac{d}{2}$ such that $p = \sum_{i=1}^{r} q_i^2$. For all $i \in \{1, \ldots, r\}$, there exists a unique $c_i \in \mathbb{R}^{\binom{n+\frac{d}{2}}{\frac{d}{2}}}$ such that $q_i(x) = c_i^\top [x]_{\frac{d}{2}}$ for all $x \in \mathbb{R}^n$.

Let $Q = \sum_{i=1}^{r} c_i c_i^\top \succeq 0$. For all $x \in \mathbb{R}^n$, we then have

$$\begin{aligned} p(x) &= \sum_{i=1}^{r} (q_i(x))^2 = \sum_{i=1}^{r} \left(c_i^\top [x]_{\frac{d}{2}}\right)^2 = \sum_{i=1}^{r} [x]_{\frac{d}{2}}^\top c_i c_i^\top [x]_{\frac{d}{2}} \\ &= [x]_{\frac{d}{2}}^\top \left(\sum_{i=1}^{r} c_i c_i^\top\right) [x]_{\frac{d}{2}} = [x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}}. \end{aligned}$$

$\Longleftarrow$) Assume that there exists a $Q \succeq 0$ such that $p(x) = [x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}}$ for all $x \in \mathbb{R}^n$. Then, there exist $c_1, \ldots, c_r \in \mathbb{R}^{\binom{n+\frac{d}{2}}{\frac{d}{2}}}$ such that $Q = \sum_{i=1}^{r} c_i c_i^\top$.

For every $i \in \{1, \ldots, r\}$, let the n-variate polynomial q_i of degree $\frac{d}{2}$ be defined by $q_i(x) = c_i^\top [x]_{\frac{d}{2}}$ for all $x \in \mathbb{R}^n$. For all $x \in \mathbb{R}^n$, we then have

$$\begin{aligned} p(x) &= [x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}} = [x]_{\frac{d}{2}}^\top \left(\sum_{i=1}^{r} c_i c_i^\top\right) [x]_{\frac{d}{2}} \\ &= \sum_{i=1}^{r} [x]_{\frac{d}{2}}^\top c_i c_i^\top [x]_{\frac{d}{2}} = \sum_{i=1}^{r} \left(c_i^\top [x]_{\frac{d}{2}}\right)^2. \end{aligned}$$

Hence, p is SOS.

□

It follows that we can determine whether a polynomial is SOS by solving a system of SDP equations (see also the proof of Theorem 3.5).

Corollary 4.1.1. An n-variate polynomial p of an even degree d is SOS if and only if there exists a $(q_{\beta,\gamma})_{\beta,\gamma \in \mathbb{N}^n_{\frac{d}{2}}} \succeq 0$ such that, for all $\alpha \in \mathbb{N}^n_d$, we have

$$p_\alpha = \sum_{\substack{\beta \in \mathbb{N}^n_{\frac{d}{2}} \\ \beta+\gamma=\alpha}} \sum_{\gamma \in \mathbb{N}^n_{\frac{d}{2}}} q_{\beta,\gamma}.$$

The set of SOS polynomials is a convex cone. That is, any nonnegative linear combination of SOS polynomials is also SOS.

Theorem 4.2. Let V be the vector space of n-variate polynomials of an even degree d, and $\mathcal{K}$ the set of SOS polynomials in V. Then, $\mathcal{K}$ is a convex cone.

Proof. Define the function $\Phi : \mathbb{S}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}}} \to V$ for all $Q \in \mathbb{S}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}}}$ by

$$\Phi(Q) = \left(x \mapsto [x]_{\frac{d}{2}}^{\top} Q [x]_{\frac{d}{2}}\right).$$

Then, Φ is a linear transformation. From Theorem 4.1, it follows that

$$\mathcal{K} = \Phi\left(\mathbb{S}_{+}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}}}\right).$$

Since $\mathbb{S}_{+}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}}}$ is a convex cone, it now follows from Theorem A.10 that $\mathcal{K}$ is also a convex cone. □

The dual cone of the set of SOS polynomials is given by the following theorem (see also Definition A.11).

Theorem 4.3. Let V be the vector space of n-variate polynomials of an even degree d, and $\mathcal{K}$ the set of SOS polynomials in V. Then, the dual cone $\mathcal{K}^*$ of $\mathcal{K}$ is given by

$$\mathcal{K}^* = \left\{ s \in V : \left(s_{\beta+\gamma}\right)_{\beta,\gamma \in \mathbb{N}^n_{\frac{d}{2}}} \succeq 0 \right\}.$$

Proof. Let $s \in V$. Then, $s^* \in \mathcal{K}$ if and only if, for all $Q = \left(q_{\beta,\gamma}\right)_{\beta,\gamma \in \mathbb{N}^n_{\frac{d}{2}}} \succeq 0$, and $p \in V$ defined by $p(x) = [x]_{\frac{d}{2}}^{\top} Q [x]_{\frac{d}{2}}$ for all $x \in \mathbb{R}^n$, we have

$$\begin{aligned}
\langle s, p \rangle &= \sum_{\alpha \in \mathbb{N}^n_d} s_\alpha p_\alpha = \sum_{\alpha \in \mathbb{N}^n_d} s_\alpha \sum_{\substack{\beta \in \mathbb{N}^n_{\frac{d}{2}} \\ \beta+\gamma=\alpha}} \sum_{\gamma \in \mathbb{N}^n_{\frac{d}{2}}} q_{\beta,\gamma} \\
&= \sum_{\beta \in \mathbb{N}^n_{\frac{d}{2}}} \sum_{\gamma \in \mathbb{N}^n_{\frac{d}{2}}} s_{\beta+\gamma} q_{\beta,\gamma} = \left\langle \left(s_{\beta+\gamma}\right)_{\beta,\gamma \in \mathbb{N}^n_{\frac{d}{2}}}, Q \right\rangle \geq 0.
\end{aligned}$$

Hence, $s \in \mathcal{K}^*$ if and only if $\left(s_{\beta+\gamma}\right)_{\beta,\gamma \in \mathbb{N}^n_{\frac{d}{2}}} \succeq 0$. □

Corollary 4.3.1. Let p be an n-variate SOS polynomial of an even degree d. For all $\left(y_\alpha\right)_{\alpha \in \mathbb{N}^n_d} \in \mathbb{R}^{\binom{n+d}{d}}$ with $\left(y_{\beta+\gamma}\right)_{\beta,\gamma \in \mathbb{N}^n_{\frac{d}{2}}} \succeq 0$, we then have

$$\sum_{\alpha \in \mathbb{N}^n_d} p_\alpha y_\alpha \geq 0.$$

4.2 SOS-convex polynomials

An n-variate polynomial p of an even degree d is convex if and only if, for all $x \in \mathbb{R}^n$, we have $\nabla^2 p(x) \succeq 0$, or $y^\top \nabla^2 p(x)\, y \geq 0$ for all $y \in \mathbb{R}^n$. In general, determining whether p is convex is NP-hard. (e.g., Ahmadi et al., 2024). However, a sufficient condition is that the polynomial $x, y \mapsto y^\top \nabla^2 p(x)\, y$ is SOS.

Definition 4.2 (SOS-convex polynomial)**.** A polynomial p of an even degree d is called *SOS-convex* if the polynomial

$$x, y \mapsto y^\top \nabla^2 p(x)\, y$$

is SOS.

Note that every SOS-convex polynomial is convex, hence the name. However, not every convex polynomial is also SOS-convex (e.g., Ahmadi et al., 2024). Nevertheless, every convex polynomial of degree two *is* also SOS-convex.

Theorem 4.4 (e.g., Ahmadi et al., 2024)**.** Every convex polynomial of degree two is also SOS-convex.

Proof. Let p be a n-variate convex polynomial of degree two. Then, there exist an $A \succeq 0$, a $b \in \mathbb{R}^n$ and a $c \in \mathbb{R}$ such that, for all $x \in \mathbb{R}^n$, we have

$$p(x) = x^\top A x + b^\top x + c.$$

Since $A \succeq 0$, there exists an $L \in \mathbb{R}^{n \times n}$ such that $A = LL^\top$. For all $x, y \in \mathbb{R}^n$, we now have

$$y^\top \nabla^2 p(x)\, y = y^\top (2A)\, y = y^\top \left(2LL^\top\right) y = 2\left(L^\top y\right)^\top L^\top y = 2\left\|L^\top y\right\|^2.$$

Hence, since the polynomial $x, y \mapsto 2\left\|L^\top y\right\|^2$ is SOS, p is SOS-convex. □

SOS-convexity is invariant with respect to affine transformations of the function variables.

Theorem 4.5. Let p be n-variate SOS-convex polynomial of an even degree d. Then, for all $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^n$, the polynomial

$$x \mapsto p(Ax + b)$$

is also SOS-convex.

Proof. Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^n$. Since p is SOS-convex, there exist $(n+n)$-variate polynomials $q_1, \ldots, q_r$ of degree $\frac{d}{2}$ such that, for all $u, v \in \mathbb{R}^n$, we have $v^\top \nabla^2 p(u)\, v = \sum_{i=1}^r (q_i(u,v))^2$. For all $x, y \in \mathbb{R}^n$, we now have

$$y^\top \left(\frac{\partial^2 p(Ax+b)}{\partial x\, \partial x^\top}\right) y = y^\top A \nabla^2 p(Ax+b)\, A^\top y = \sum_{i=1}^r \left(q_i\left(Ax+b, A^\top y\right)\right)^2,$$

so the polynomial $x, y \mapsto y^\top \left(\frac{\partial^2 p(Ax+b)}{\partial x\, \partial x^\top}\right) y$ is SOS. Hence, the polynomial $x \mapsto p(Ax+b)$ is SOS-convex. $\square$

Determining whether a polynomial is SOS-convex is similar to determining whether a polynomial is SOS.

Let p be an n-variate polynomial of an even degree d. From the power rule, it then follows that, for all $x \in \mathbb{R}^n$ and $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, we have

$$\begin{aligned}
y^\top \nabla^2 p(x)\, y &= \sum_{i=1}^n \sum_{j=1}^n \left(\frac{\partial^2}{\partial x^{e_i+e_j}} \sum_{\alpha \in \mathbb{N}_d^n} p_\alpha x^\alpha\right) y_i y_j \\
&= \sum_{i=1}^n \sum_{j=1}^n \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ \alpha \geq e_i + e_j}} \frac{\alpha!}{(\alpha - e_i - e_j)!} p_\alpha x^{\alpha - e_i - e_j} y_i y_j \\
&= \sum_{i=1}^n \sum_{j=1}^n \sum_{\beta \in \mathbb{N}_{d-2}^n} \frac{(\beta + e_i + e_j)!}{\beta!} p_{\beta + e_i + e_j} x^\beta y_i y_j.
\end{aligned}$$

The following theorem and corollary now follow directly from Theorem 4.1 and Corollary 4.1.1.

Theorem 4.6 (e.g., Ahmadi et al., 2024)**.** An n-variate polynomial p of an even degree d is SOS-convex if and only if there exists a $Q \succeq 0$ such that, for all $x, y \in \mathbb{R}^n$, we have

$$y^\top \nabla^2 p(x)\, y = \left([x]_{\frac{d}{2}-1} \otimes y\right)^\top Q \left([x]_{\frac{d}{2}-1} \otimes y\right).$$

Corollary 4.6.1. An n-variate polynomial p of an even degree d is SOS-convex if and only if there exists a

$$\left(q_{(\beta,i),(\gamma,j)}\right)_{(\beta,i),(\gamma,j) \in \mathbb{N}_{\frac{d}{2}-1}^n \times \{1,\ldots,n\}} \succeq 0$$

such that, for all $\alpha \in \mathbb{N}_{d-2}^n$ and $i, j \in \{1, \ldots, n\}$, we have

$$\frac{(\alpha + e_i + e_j)!}{\alpha!} p_{\alpha + e_i + e_j} = \sum_{\substack{\beta \in \mathbb{N}_{\frac{d}{2}-1}^n \\ \beta + \gamma = \alpha}} \sum_{\gamma \in \mathbb{N}_{\frac{d}{2}-1}^n} q_{(\beta,i),(\gamma,j)}.$$

Similarly to SOS polynomials, the set of SOS-convex polynomials is a convex cone.

Theorem 4.7. Let V be the vector space of n-variate polynomials of an even degree d, and $\tilde{\mathcal{K}}$ the set of SOS-convex polynomials in V. Then, $\tilde{\mathcal{K}}$ is a convex cone.

Proof. Let V' be the vector space of $(n+n)$-variate polynomials of an even degree d, and $\mathcal{K}'$ the cone of SOS polynomials in V'. Define the function $H : V \mapsto V'$ for all $p \in V$ by

$$H(p) = \left(x, y \mapsto y^\top \nabla^2 p(x)\, y\right).$$

Then, H is a linear transformation. From Theorem 4.6, it follows that

$$\tilde{\mathcal{K}} = \{p \in V : H(p) \in \mathcal{K}'\}.$$

Since $\mathcal{K}'$ is a convex cone, it now follows from Theorem A.11 that $\tilde{\mathcal{K}}$ is also a convex cone. □

4.3 Jensen-type inequality

In this section, we show that SOS-convex polynomials satisfy a Jensen-type inequality. First, we need the following theorem.

Theorem 4.8 (e.g., Helton and Nie, 2010)**.** Let p be an n-variate SOS-convex polynomial of an even degree d, and let $x \in \mathbb{R}^n$. Then, the n-variate polynomial q of degree d, defined for all $y \in \mathbb{R}^n$ by

$$q(y) = p(y) - p(x) - (\nabla p(x))^\top (y - x),$$

is SOS.

Proof. From the chain rule and the fundamental theorem of calculus (see also the proof of the descent lemma in Appendix B), it then follows that, for all $y \in \mathbb{R}^n$, we have

$$\begin{aligned} q(y) &= \int_0^1 (\nabla p(x + t(y-x)) - \nabla p(x))^\top (y-x)\, dt \\ &= \int_0^1 \left(\int_0^t \nabla^2 p(x + s(y-x))(y-x)\, ds\right)^\top (y-x)\, dt \\ &= \int_0^1 \int_0^t (y-x)^\top \nabla^2 p(x + s(y-x))(y-x)\, ds\, dt. \end{aligned}$$

Since p is SOS-convex, there exists a $Q \succeq 0$ such that, for all $y \in \mathbb{R}^n$, we have

$$q(y) = \int_0^1 \int_0^t \left([x + s(y-x)]_{\frac{d}{2}-1} \otimes (y-x)\right)^\top Q \left([x + s(y-x)]_{\frac{d}{2}-1} \otimes (y-x)\right) ds\, dt.$$

From the multi-binomial theorem, it follows that, for all $\beta \in \mathbb{N}^n_{\frac{d}{2}-1}$ and $i \in \{1, \ldots, n\}$, we have

$$\begin{aligned}
(x + s(y-x))^\beta (y_i - x_i) &= \left(\sum_{\substack{\nu \in \mathbb{N}_0^n \\ \nu \leq \beta}} \binom{\beta}{\nu} x^\nu (s(y-x))^{\beta-\nu} \right) (y_i - x_i) \\
&= \sum_{\substack{\nu \in \mathbb{N}_0^n \\ \nu \leq \beta}} \binom{\beta}{\nu} s^{|\beta|-|\nu|} x^\nu (y-x)^{\beta-\nu+e_i} \\
&= \sum_{\substack{\gamma \in \mathbb{N}_0^n \\ e_i \leq \gamma \leq \beta + e_i}} \binom{\beta}{\gamma - e_i} s^{|\gamma|-1} x^{\beta-\gamma+e_i} (y-x)^\gamma
\end{aligned}$$

for all $y \in \mathbb{R}^n$ and $s \in [0,1]$.

Define the function $C : [0,1] \to \mathbb{R}^{\binom{n+\frac{d}{2}-1}{\frac{d}{2}-1} n \times \binom{n+\frac{d}{2}}{\frac{d}{2}}}$ for all $s \in [0,1]$ by

$$C(s) = \left([e_i \leq \gamma \leq \beta + e_i] \binom{\beta}{\gamma - e_i} s^{|\gamma|-1} x^{\beta-\gamma+e_i} \right)_{(\beta,i) \in \mathbb{N}^n_{\frac{d}{2}-1} \times \{1,\ldots,n\}, \gamma \in \mathbb{N}^n_{\frac{d}{2}}}.$$

For all $y \in \mathbb{R}^n$, we then have

$$q(y) = [y-x]_{\frac{d}{2}}^\top \left(\int_0^1 \int_0^t C(s) Q (C(s))^\top \, ds \, dt \right) [y-x]_{\frac{d}{2}},$$

and

$$\int_0^1 \int_0^t C(s) Q (C(s))^\top \, ds \, dt \succeq 0.$$

Hence, q is SOS. $\square$

It follows that, if an SOS-convex polynomial p that is bounded from below with lower bound m attains its minimum, then the polynomial $p - m$ is SOS.

Corollary 4.8.1 (e.g., Helton and Nie, 2010)**.** Let p be an SOS-convex polynomial that is bounded from below with lower bound m. Assume that p attains its minimum. Then, the polynomial $p - m$ is SOS.

Proof. Let x^* be a minimizer of p. Then, $p(x^*) \geq m$. Since $\nabla p(x^*) = 0$, it follows that the polynomial

$$x \mapsto p(x) - p(x^*) - (\nabla p(x^*))^\top (x - x^*) = p(x) - p(x^*)$$

is SOS. Hence, the polynomial

$$p - m = p - p(x^*) + (p(x^*) - m)$$

is also SOS. $\square$

We can now prove the Jensen-type inequality.

Theorem 4.9 (Jensen-type inequality; e.g., Lasserre, 2009)**.** Let p be an n-variate SOS-convex polynomial of an even degree d. For all $(y_\alpha)_{\alpha\in\mathbb{N}_d^n} \in \mathbb{R}^{\binom{n+d}{d}}$ with $y_0 = 1$ and $(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} \succeq 0$, and $x = (x_1,\ldots,x_n) = (y_{e_1},\ldots,y_{e_n})$, we then have

$$\sum_{\alpha\in\mathbb{N}_d^n} p_\alpha x^\alpha \leq \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha y_\alpha.$$

Proof. Let $(y_\alpha)_{\alpha\in\mathbb{N}_d^n} \in \mathbb{R}^{\binom{n+d}{d}}$ with $y_0 = 1$ and $(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} \succeq 0$, and let $x = (x_1,\ldots,x_n) = (y_{e_1},\ldots,y_{e_n})$. Since p is SOS-convex, the n-variate polynomial q of degree d, defined by $q(z) = p(z) - p(x) - (\nabla p(x))^\top (z-x)$ for all $z\in\mathbb{R}^n$, is SOS. For all $z = (z_1,\ldots,z_n)\in\mathbb{R}^n$, we have

$$\sum_{\alpha\in\mathbb{N}_d^n} q_\alpha z^\alpha = \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha z^\alpha - \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha x^\alpha - \sum_{i=1}^n \frac{\partial p(x_1,\ldots,x_n)}{\partial x_i}(z_i - x_i).$$

Since q is SOS, $y_0 = 1$ and $(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} \succeq 0$, it now follows from Corollary 4.3.1 that

$$\begin{aligned}\sum_{\alpha\in\mathbb{N}_d^n} q_\alpha y_\alpha &= \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha y_\alpha - \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha x^\alpha y_0 - \sum_{i=1}^n \frac{\partial p(x_1,\ldots,x_n)}{\partial x_i}(y_{e_i} - x_i).\\ &= \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha y_\alpha - \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha x^\alpha \geq 0.\end{aligned}$$

Hence, $\sum_{\alpha\in\mathbb{N}_d^n} p_\alpha x^\alpha \leq \sum_{\alpha\in\mathbb{N}_d^n} p_\alpha y_\alpha$. □

Chapter 5

Polynomial optimization

5.1 Polynomial optimization problems

In this chapter, we analyze a particular class of mathemaical optimization problems called *polynomial* optimization problems. A polynomial optimization problem is an optimization problem with a polynomial objective function and polynomial constraints.

Definition 5.1 (Polynomial optimization problem)**.** An *n-variate polynomial optimization problem of degree d* is an optimization problem

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \quad & \sum_{\alpha \in \mathbb{N}^n_d} f_\alpha x^\alpha \\ \text{s.t.} \quad & \sum_{\alpha \in \mathbb{N}^n_d} g_{i,\alpha} x^\alpha \leq 0, \quad i \in \{1, \ldots, m\} \end{aligned} \tag{P}$$

where f and $g_1, \ldots, g_m$ are n-variate polynomials of degree d.

In general, a polynomial optimization problem is NP-hard (Bellare and Rogaway, 1995).

We assume that d is even. Then, we can reformulate (P) as an SDP with a rank constraint.

Theorem 5.1 (e.g., Lasserre, 2009)**.** The optimization problem

$$\begin{aligned}
\inf_{(y_\alpha)_{\alpha\in\mathbb{N}^n_d}\in\mathbb{R}^{\binom{n+d}{d}}} \quad & \sum_{\alpha\in\mathbb{N}^n_d} f_\alpha y_\alpha \\
\text{s. t.} \quad & \sum_{\alpha\in\mathbb{N}^n_d} g_{i,\alpha} y_\alpha \leq 0, \qquad\qquad i\in\{1,\ldots,m\} \\
& y_0 = 1 \\
& (y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} \succeq 0 \\
& \operatorname{rank}(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} = 1
\end{aligned}$$

is equivalent with problem (P).

Proof. $\Longrightarrow$) Let $x\in\mathbb{R}^n$ and $(y_\alpha)_{\alpha\in\mathbb{N}^n_d} = [x]_d$. Then, $y_0 = x^0 = 1$. Moreover, for all $\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}$, we have $y_{\beta+\gamma} = x^{\beta+\gamma} = x^\beta\cdot x^\gamma$. Hence, $(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} = [x]_{\frac{d}{2}}[x]_{\frac{d}{2}}^\top \succeq 0$, and $\operatorname{rank}(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} = 1$.

$\Longleftarrow$) Let $(y_\alpha)_{\alpha\in\mathbb{N}^n_d}\in\mathbb{R}^{\binom{n+d}{d}}$ with $y_0 = 1$, $(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}}\succeq 0$, and

$\operatorname{rank}(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} = 1$. Then, there exists a $z = (z_\beta)_{\beta\in\mathbb{N}^n_{\frac{d}{2}}}\in\mathbb{R}^{\binom{n+\frac{d}{2}}{\frac{d}{2}}}$ such that $(y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} = zz^\top$, or $y_{\beta+\gamma} = z_\beta\cdot z_\gamma$ for all $\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}$. For all $\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}$, we now have

$$\begin{aligned}
y_{\beta+\gamma} &= z_\beta\cdot z_\gamma = z_\beta\cdot z_\gamma\cdot 1 = z_\beta\cdot z_\gamma\cdot y_0 = z_\beta\cdot z_\gamma\cdot y_{0+0} \\
&= z_\beta\cdot z_\gamma\cdot z_0\cdot z_0 = z_\beta\cdot z_0\cdot z_\gamma\cdot z_0 = y_{\beta+0}\cdot y_{\gamma+0} = y_\beta\cdot y_\gamma.
\end{aligned}$$

Let $x = (x_1,\ldots,x_n) = (y_{e_1},\ldots,y_{e_n})$. For all $\alpha = (\alpha_1,\ldots,\alpha_n)\in\mathbb{N}^n_d$, we then have

$$y_\alpha = y_{\alpha_1 e_1+\cdots+\alpha_n e_n} = y_{e_1}^{\alpha_1}\cdots y_{e_n}^{\alpha_n} = x_1^{\alpha_1}\cdots x_n^{\alpha_n} = x^\alpha.$$

Hence, $(y_\alpha)_{\alpha\in\mathbb{N}^n_d} = [x]_d$.

$\square$

Unfortunately, an SDP does not allow a rank constraint. However, if we relax the rank constraint, we obtain a relaxation of problem (P).

Corollary 5.1.1 (e.g., Lasserre, 2009)**.** The optimization problem

$$\begin{aligned}
\inf_{(y_\alpha)_{\alpha\in\mathbb{N}^n_d}\in\mathbb{R}^{\binom{n+d}{d}}} \quad & \sum_{\alpha\in\mathbb{N}^n_d} f_\alpha y_\alpha \\
\text{s.t.} \quad & \sum_{\alpha\in\mathbb{N}^n_d} g_{i,\alpha} y_\alpha \le 0, \qquad i\in\{1,\ldots,m\} \\
& y_0 = 1 \\
& (y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} \succeq 0
\end{aligned} \tag{SDP}$$

is an SDP relaxation of (P).

The dual problem of (SDP) is called the *first-order Lasserre relaxation* of (P) (e.g., Lasserre, 2009), and is given by the following theorem (see also Definition A.15).

Theorem 5.2 (e.g., Lasserre, 2009)**.** The dual problem of (SDP) is given by

$$\begin{aligned}
\sup_{\substack{t\in\mathbb{R}\\ \lambda_1,\ldots,\lambda_m\in\mathbb{R}}} \quad & t \\
\text{s.t.} \quad & f - t + \sum_{i=1}^m \lambda_i g_i \text{ is SOS} \\
& \lambda_i \ge 0, \quad i\in\{1,\ldots,m\}\,.
\end{aligned} \tag{Lasserre}$$

Proof. Let $\mathcal{K} = \left\{(y_\alpha)_{\alpha\in\mathbb{N}^n_d}\in\mathbb{R}^{\binom{n+d}{d}} : (y_{\beta+\gamma})_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} \succeq 0\right\}$. The Lagrangian L of (SDP) is then given for all $(y_\alpha)_{\alpha\in\mathbb{N}^n_d}\in\mathcal{K}$, $\lambda_1,\ldots,\lambda_m\ge 0$, and $t\in\mathbb{R}$ by

$$\begin{aligned}
L\left((y_\alpha)_{\alpha\in\mathbb{N}^n_d}, \lambda_1,\ldots,\lambda_m, t\right) &= \sum_{\alpha\in\mathbb{N}^n_d} f_\alpha y_\alpha + \sum_{i=1}^m \lambda_i \sum_{\alpha\in\mathbb{N}^n_d} g_{i,\alpha} y_\alpha + t\,(1-y_0) \\
&= t + \sum_{\alpha\in\mathbb{N}^n_d}\left(f_\alpha - [\alpha=0]\cdot t + \sum_{i=1}^m \lambda_i g_{i,\alpha}\right) y_\alpha
\end{aligned}$$

From Corollary 4.3.1, it follows that, for all $\lambda_1,\ldots,\lambda_m\ge 0$ and $t\in\mathbb{R}$, we have

$$\inf\left\{L\left((y_\alpha)_{\alpha\in\mathbb{N}^n_d}, \lambda_1,\ldots,\lambda_m, t\right) : (y_\alpha)_{\alpha\in\mathbb{N}^n_d}\in\mathcal{K}\right\} = \begin{cases} t & \text{if } f - t + \sum_{i=1}^m \lambda_i g_i \text{ is SOS,} \\ -\infty & \text{otherwise.}\end{cases}$$

Hence, the dual problem of (SDP) is given by (Lasserre). □

Recall that we can determine whether a polynomial is SOS by solving a system of SDP equations. Therefore, (Lasserre) can be reformulated as another SDP.

Let $(\mathrm{P})^*$, $(\mathrm{SDP})^*$, and $(\mathrm{Lasserre})^*$ be the optimal solution values of problems (P), (SDP), and (Lasserre) respectively. Since (SDP) is a relaxation of (P), we then have $(\mathrm{SDP})^* \leq (\mathrm{P})^*$. From weak duality (see also Corollary A.15.1), it follows that $(\mathrm{Lasserre})^* \leq (\mathrm{SDP})^*$ Hence,

$$(\mathrm{Lasserre})^* \leq (\mathrm{SDP})^* \leq (\mathrm{P})^* ,$$

so (Lasserre) is indeed a relaxation of problem (P).

Lasserre (2009) and De Klerk and Laurent (2011) provide sufficient conditions under which the first-order Lasserre relaxation is exact, i.e., $(\mathrm{Lasserre})^* = (\mathrm{P})^*$. In that case,

$$(\mathrm{Lasserre})^* = (\mathrm{SDP})^* = (\mathrm{P})^* ,$$

so the SDP relaxation is also exact. We can then solve (P) by solving (SDP).

5.2 SOS-convex polynomial optimization

A sufficient condition under which we can solve (P) by solving (SDP) is that (P) is an *SOS-convex* polynomial optimization problem.

Definition 5.2 (SOS-convex polynomial optimization problem)**.** An *n-variate SOS-convex polynomial optimization problem* of an even degree d is an optimization problem

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \quad & \sum_{\alpha \in \mathbb{N}^n_d} f_\alpha x^\alpha \\ \text{s.t.} \quad & \sum_{\alpha \in \mathbb{N}^n_d} g_{i,\alpha} x^\alpha \leq 0, \quad i \in \{1, \ldots, m\} \end{aligned} \tag{SOS-CVX}$$

where f and $g_1, \ldots, g_m$ are n-variate SOS-convex polynomials of an even degree d.

Theorem 5.3. The optimization problem

$$\begin{aligned}
\inf_{(y_\alpha)_{\alpha\in\mathbb{N}^n_d}\in\mathbb{R}^{\binom{n+d}{d}}} \quad & \sum_{\alpha\in\mathbb{N}^n_d} f_\alpha y_\alpha \\
\text{s.t.} \quad & \sum_{\alpha\in\mathbb{N}^n_d} g_{i,\alpha} y_\alpha \le 0, \qquad i\in\{1,\dots,m\} \\
& y_0 = 1 \\
& \left(y_{\beta+\gamma}\right)_{\beta,\gamma\in\mathbb{N}^n_{\frac{d}{2}}} \succeq 0
\end{aligned} \tag{SOS-CVX-SDP}$$

is an SDP relaxation of (SOS-CVX).

An optimal solution to (SOS-CVX-SDP) now yields an optimal solution to (SOS-CVX).

Theorem 5.4 (e.g., Lasserre, 2009)**.** Assume that (SOS-CVX-SDP) has an optimal solution $(y^*_\alpha)_{\alpha\in\mathbb{N}^n_d}$. Then,

$$\text{(SOS-CVX-SDP)}^* = \text{(SOS-CVX)}^*,$$

and $x^* = (x^*_1,\dots,x^*_n) = \left(y^*_{e_1},\dots,y^*_{e_n}\right)$ is an optimal solution to (SOS-CVX).

Proof. Since $g_1,\dots,g_m$ are SOS-convex, it follows from the Jensen-type equality that, for all $i\in\{1,\dots,m\}$, we have

$$\sum_{\alpha\in\mathbb{N}^n_d} g_{i,\alpha} x^{*\alpha} \le \sum_{\alpha\in\mathbb{N}^n_d} g_{i,\alpha} y^*_\alpha \le 0.$$

Therefore, x^* is a feasible solution to (SOS-CVX). Since f is also SOS-convex, it follows from the Jensen-type inequality that

$$\text{(SOS-CVX-SDP)}^* \le \text{(SOS-CVX)}^* \le \sum_{\alpha\in\mathbb{N}^n_d} f_\alpha x^{*\alpha} \le \sum_{\alpha\in\mathbb{N}^n_d} f_\alpha y^*_\alpha = \text{(SOS-CVX-SDP)}^*.$$

Hence, (SOS-CVX-SDP)* = (SOS-CVX)*, and x^* is an optimal solution to (SOS-CVX). □

Chapter 6

Algebraically open sets and conic decompositions

6.1 Algebraically open sets

Let V be a finite-dimensional normed vector space, and $A \subseteq V$. Then, a point $x \in A$ is an *interior point* of A if a ball around x is contained in A, and A is an *open* set if every point $x \in A$ is an interior point of A.

Definition 6.1 (Open set)**.** Let V be an finite-dimensional normed vector space, and $A \subseteq V$. A point $x \in A$ is then called an *interior point* of A if there exists an $\varepsilon > 0$ such that, for all $y \in V$ with $\|y - x\| < \varepsilon$, we have $y \in A$.

The set of interior points of A is denoted by $\operatorname{int}(A)$. A is called an *open set* if $A = \operatorname{int}(A)$.

We can generalize the notion of an open set to all vector spaces using the notion of an *algebraically open* set.

Let V be a vector space, and $A \subseteq V$. Then, a point $x \in A$ is an *algebraic interior point* of A if x lies in the interior of every intersection of A with a line through x, and A is an *algebraically open set* if every point $x \in A$ is an algebraic interior point of A.

Definition 6.2 (Algebraically open set)**.** Let V be a vector space, and $A \subseteq V$. A point $x \in A$ is then called an *algebraic interior point* of A if, for all $u \in V \setminus \{0\}$, there exists an $\varepsilon > 0$ such that, for all $t \in \mathbb{R}$ with $|t| < \varepsilon$, we have $x + tu \in A$.

The set of algebraic interior points of A is denoted by $\operatorname{aint}(A)$. A is called an *algebraically open set* if $A = \operatorname{aint}(A)$.

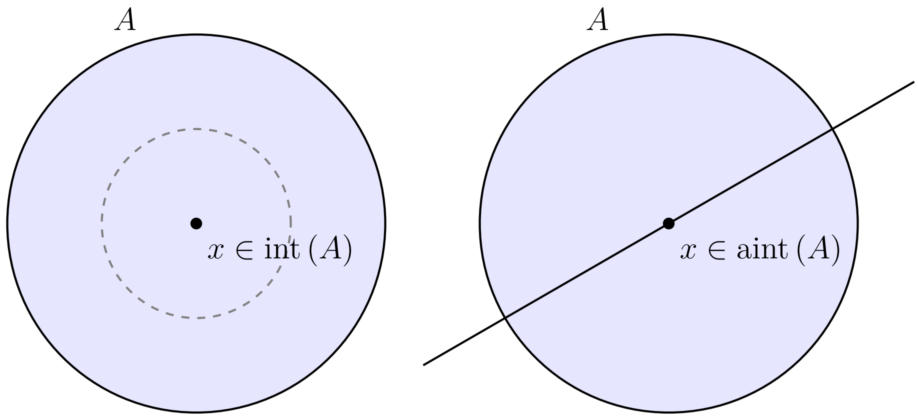

Figure 6.1: $x \in \mathrm{int}\,(A)$ if a ball around x is contained in A, and $x \in \mathrm{aint}\,(A)$ if x lies in the interior of every intersection of A with a line through x.

An equivalent definition of an algebraically open set A is that every intersection of A with a line through A is an open set.

Theorem 6.1 (e.g., Barvinok, 2002, Definition II.1.5)**.** Let V be a vector space, and $A \subseteq V$. Then, A is an algebraically open set if and only if, for all $x \in A$ and $u \in V \setminus \{0\}$, the set

$$\{t \in \mathbb{R} : x + tu \in A\}$$

is open.

Proof. $\Longleftarrow$) Assume that $\{t \in \mathbb{R} : x + tu \in A\}$ is an open set for all $x \in A$ and $u \in V \setminus \{0\}$. Let $x \in A$, $u \in V \setminus \{0\}$, and $I = \{t \in \mathbb{R} : x + tu \in A\}$. Then, $0 \in I$, and I is an open set, so $0 \in \mathrm{int}\,(I)$. Therefore, there exists an $\varepsilon > 0$ such that, for all $t \in \mathbb{R}$ with $|t| < \varepsilon$, we have $t \in I$, or $x + tu \in A$. Therefore, $x \in \mathrm{aint}\,(A)$. Hence, A is an algebraically open set.

$\Longrightarrow$) Assume that A is an algebraically open set. Suppose that there exist an $x \in A$ and a $u \in V \setminus \{0\}$, such that the set $I = \{t \in \mathbb{R} : x + tu \in A\}$ is not open. Then, there exists a $t \in I$ with $t \notin \mathrm{int}\,(I)$. Since $t \in I$, we have $x + tu \in A$. Since $t \notin \mathrm{int}\,(I)$, it follows that, for all $\varepsilon > 0$, there exists an $s \in \mathbb{R}$ with $|s - t| < \varepsilon$, such that $s \notin I$, or

$$x + su = x + tu + (s - t)\,u \notin A.$$

Therefore, $x + tu \notin \mathrm{aint}\,(A)$. However, A would then not be an algebraically open set. Hence, $\{t \in \mathbb{R} : x + tu \in A\}$ is an open set for all $x \in A$ and $u \in V \setminus \{0\}$.

□

If V is a finite-dimensional normed vector space, then every interior point of A is also an algebraic interior point of A.

Theorem 6.2. Let V be a finite-dimensional normed vector space, and $A \subseteq V$. Then,
$$\mathrm{int}\,(A) \subseteq \mathrm{aint}\,(A)\,.$$

Proof. Let $x \in \mathrm{int}\,(A)$. Then, there exists an $\varepsilon > 0$ such that, for all $y \in \mathbb{R}^n$ with $\|y - x\| < \varepsilon$, we have $y \in A$. Let $u \in V \setminus \{0\}$. Then, for all $t \in \mathbb{R}$ with $|t| < \frac{\varepsilon}{\|u\|}$, we have
$$\|(x + tu) - x\| = \|tu\| = |t|\,\|u\| < \varepsilon,$$
so $x + tu \in A$. Hence, $x \in \mathrm{aint}\,(A)$. □

Moreover, if A is a convex set, then the interior and algebraic interior of A coincide.

Theorem 6.3. Let V be an finite-dimensional normed vector space, and $A \subseteq V$ a convex set. Then,
$$\mathrm{aint}\,(A) = \mathrm{int}\,(A)\,.$$

Proof. $\supseteq$) Follows directly from Theorem 6.2.

$\subseteq$) Let $x \in \mathrm{aint}\,(A)$, and $\{u_1, \ldots, u_n\}$ an orthonormal basis of V (see also Theorem A.6 and Theorem A.7). Then, for every $i \in \{1, \ldots, n\}$, there exists an $\varepsilon_i > 0$ such that, for all $t_i \in \mathbb{R}$ with $|t_i| < \varepsilon_i$, we have $x + t_i u_i \in A$.

Let $y \in \mathbb{R}^n$ with $\|y - x\| < \frac{\min\{\varepsilon_1, \ldots, \varepsilon_n\}}{n}$, and $t_1, \ldots, t_n \in \mathbb{R}$ such that $y - x = t_1 u_1 + \cdots + t_n u_n$. From Corollary A.8.1, it then follows that, for all $i \in \{1, \ldots, n\}$, we have
$$|nt_i| = n\,|t_i| \leq n\,\|y - x\| < \min\{\varepsilon_1, \ldots, \varepsilon_n\} \leq \varepsilon_i,$$
so $x + nt_i u_i \in A$. Therefore, since A is a convex set, we have
$$y = x + \sum_{i=1}^{n} t_i u_i = \frac{1}{n} \sum_{i=1}^{n} (x + nt_i u_i) \in A.$$

Hence, $x \in \mathrm{int}\,(A)$.

□

The requirement that A needs to be a convex set in Theorem 6.3 is necessary, as shown by the following example.

Example 6.1. Let $A = \{(x,y) \in \mathbb{R}^2 : |y| \geq x^2\} \cup \{(x,0) : x \in \mathbb{R}\}$. Then, A is not a convex set. We can now show that $(0,0) \in \operatorname{aint}(A)$, but $(0,0) \notin \operatorname{int}(A)$.

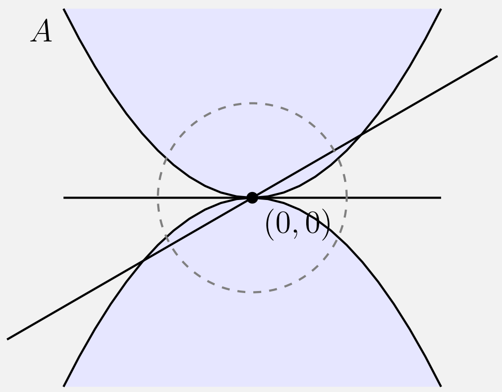

Figure 6.2: $(0,0) \in \operatorname{aint}(A)$, but $(0,0) \notin \operatorname{int}(A)$.

Let $(u,v) \in \mathbb{R}^2$. If $u = 0$, then $(tu, tv) = (0, tv) \in A$ for all $t \in \mathbb{R}$. If $v = 0$, then $(tu, tv) = (tu, 0) \in A$ for all $t \in \mathbb{R}$. If $u \neq 0$ and $v \neq 0$, then, for all $t \in \mathbb{R}$ with $|t| < \frac{|v|}{u^2}$, we have $t^2 \leq \frac{|t|v}{u^2}$, so

$$|tv| = |t|\,|v| \geq t^2 u^2 = (tu)^2\,,$$

so $(tu, tv) \in A$. Hence, $(0,0) \in \operatorname{aint}(A)$.

Let $\varepsilon > 0$, $t = \min\left\{\frac{\varepsilon}{\sqrt{5}}, 1\right\}$ and $(x,y) = \left(t, \frac{t^2}{2}\right)$. Then,

$$\|(x,y)\| = \sqrt{t^2 + \frac{t^4}{4}} \leq \sqrt{\frac{5t^2}{4}} = \frac{\sqrt{5}t}{2} \leq \frac{\varepsilon}{2} < \varepsilon,$$

but $(x,y) \notin A$. Hence, $(0,0) \notin \operatorname{int}(A)$.

6.2 Conic decompositions

Let V be a vector space, and $A, B \subseteq V$. Then, the set of points in V that can be written as the sum of a point in A and a point in B is called the *Minkowski sum* of A and B. Similarly, the set of points in V that can be written as the difference of a point A and a point B is called the *Minkowski difference* of A and B.

Definition 6.3 (Minkowski sum and difference)**.** Let V be a vector space, and let $A, B \subseteq V$. Then, the *Minkowski sum* $A + B$ of A and B is defined as

$$A + B = \{a + b : a \in A, b \in B\},$$

and the *Minkowski difference* $A - B$ of A and B is defined as

$$A - B = \{a - b : a \in A, b \in B\}.$$

Let V be a vector space, and $\mathcal{K}$ a cone in V. If $\operatorname{aint}(\mathcal{K}) \neq \emptyset$, then every point in V can be written as the difference of two points in $\mathcal{K}$.

Theorem 6.4 (e.g., Ahmadi and Hall, 2018)**.** Let V be a vector space, and $\mathcal{K}$ a cone in V. Assume that $\operatorname{aint}(\mathcal{K}) \neq \emptyset$. Then,

$$V = \mathcal{K} - \mathcal{K}.$$

Proof. $\supseteq$) Trivial.

$\subseteq$) Let $v \in V$, and $k \in \operatorname{aint}(\mathcal{K})$. Then, there exists an $\varepsilon > 0$ such that, for all $t \in \mathbb{R}$ with $|t| < \varepsilon$, we have $k + tv \in \mathcal{K}$. Since $k + \frac{\varepsilon}{2}v \in \mathcal{K}$ and $k \in \mathcal{K}$, and $\mathcal{K}$ is a cone, we have $\frac{2}{\varepsilon}\left(k + \frac{\varepsilon}{2}v\right) \in \mathcal{K}$ and $\frac{2}{\varepsilon}k \in \mathcal{K}$. Hence,

$$v = \left(\frac{2}{\varepsilon}k + v\right) - \frac{2}{\varepsilon}k = \frac{2}{\varepsilon}\left(k + \frac{\varepsilon}{2}v\right) - \frac{2}{\varepsilon}k \in \mathcal{K} - \mathcal{K}.$$

$\square$

The requirement that $\operatorname{aint}(\mathcal{K}) \neq \emptyset$ in Theorem 6.4 is necessary. For example, the set $\mathcal{K} = \{(x, 0) : x \in \mathbb{R}\}$ is a cone in $\mathbb{R}^2$, but $\operatorname{aint}(\mathcal{K}) = \emptyset$, and

$$\mathcal{K} - \mathcal{K} = \{(x, 0) : x \in \mathbb{R}\} \subsetneq \mathbb{R}^2.$$

6.3 The cone of SOS-convex polynomials

Let V be the vector space of n-variate polynomials of an even degree d, and $\tilde{\mathcal{K}}$ the convex cone of SOS-convex polynomials in V. Since V is a finite-dimensional vector space, and $\tilde{\mathcal{K}}$ a convex set, we then have

$$\operatorname{aint}\left(\tilde{\mathcal{K}}\right) = \operatorname{int}\left(\tilde{\mathcal{K}}\right).$$

Let $p \in V$. Recall that $p \in \tilde{\mathcal{K}}$ if and only if there exists a $Q \succeq 0$ such that $y^\top \nabla^2 p(x)\, y = \left([x]_{\frac{d}{2}-1} \otimes y\right)^\top Q \left([x]_{\frac{d}{2}-1} \otimes y\right)$ for all $x, y \in \mathbb{R}^n$. If we can take $Q \succ 0$, then $p \in \operatorname{int}\left(\tilde{\mathcal{K}}\right)$.

Theorem 6.5 (e.g., Ahmadi et al., 2024)**.** Let V be the vector space of n-variate polynomials of an even degree d, and $\tilde{\mathcal{K}}$ the cone of SOS-convex polynomials in V. Let $p \in V$. Assume that there exists a $Q \succ 0$ such that, for all $x, y \in \mathbb{R}^n$, we have

$$y^\top \nabla^2 p(x) = \left([x]_{\frac{d}{2}-1} \otimes y\right)^\top Q \left([x]_{\frac{d}{2}-1} \otimes y\right).$$

Then, $p \in \operatorname{int}\left(\tilde{\mathcal{K}}\right)$.

Proof. Let $q \in V$. Then, there exists a $Q' \in \mathbb{S}^{\binom{n+\frac{d}{2}-1}{\frac{d}{2}-1}n \times \binom{n+\frac{d}{2}-1}{\frac{d}{2}-1}n}$ such that, for all $x, y \in \mathbb{R}^n$, we have

$$y^\top \nabla^2 q(x)\, y = \left([x]_{\frac{d}{2}-1} \otimes y\right)^\top Q' \left([x]_{\frac{d}{2}-1} \otimes y\right).$$

Since $Q \succ 0$, there exists an $\varepsilon > 0$ such that, for all $t \in \mathbb{R}$ with $|t| < \varepsilon$, we have $Q + tQ' \succeq 0$, so $p + tq \in \tilde{\mathcal{K}}$. Hence, $p \in \operatorname{aint}\left(\tilde{\mathcal{K}}\right) = \operatorname{int}\left(\tilde{\mathcal{K}}\right)$. $\square$

Corollary 6.5.1. Let V be the vector space of n-variate polynomials of an even degree d, and $\tilde{\mathcal{K}}$ the cone of SOS-convex polynomials in V. Then,

$$\operatorname{int}\left(\tilde{\mathcal{K}}\right) \neq \emptyset.$$

It now follows directly from Theorem 6.4 that every polynomial can be written as the difference of two SOS-convex polynomials.

Corollary 6.5.2 (e.g., Ahmadi and Hall, 2018)**.** For every polynomial p, there exist two SOS-convex polynomials q_1 and q_2 such that

$$p = q_2 - q_1.$$

Chapter 7

$d^{\mathbf{th}}$-order Newton method

Newton's method aims to minimize a function by iteratively minimizing second-order Taylor approximations. Therefore, a natural idea to improve upon Newton's method is to minimize Taylor approximations of a higher order $d \geq 3$. Ahmadi et al. (2024) developed the *d^{th}-order Newton method*, which aims to minimize a function by iteratively minimizing regularized d^{th}-order Taylor approximations.

7.1 Regularized Taylor approximations

Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}$ is a d times continuously differentiable function on an open domain $D \subseteq \mathbb{R}^n$. Similarly to gradient descent and the classical Newton method, the d^{th}-order Newton method starts from an initial point $x_0 \in D$ and aims to construct a sequence $(x_k)_{k=0}^{\infty}$ that converges to a minimizer x^* of f.

A natural idea to update the current iterate x_k is to minimize the d^{th}-order Taylor approximation $T_{x_k,d}$ of f at x_k. However, there are some issues with this approach (Ahmadi et al., 2024):

- $T_{x_k,d}$ might be unbounded from below.
- Even if $T_{x_k,d}$ is bounded from below, it might not attain its minimum.
- A minimizer of $T_{x_k,d}$ might not be unique.
- In general, finding a minimizer of $T_{x_k,d}$ is NP-hard.

Therefore, Ahmadi et al. (2024) instead consider a regularized d^{th}-order Taylor approximation of f at x_k, with penalty factor $t \geq 0$, given by

$$x \mapsto T_{x_k,d}(x) + t \left\|x - x_k\right\|^{d'},$$

where d' is the smallest even number greater than d, ensuring that this function is a polynomial. The regularization term imposes a penalty on moving away from x_k, where $T_{x_k,d}$ becomes less accurate. Therefore, minimizing the regularized Taylor approximation is a more conservative approach than minimizing the Taylor approximation itself.

An important question is how large the penalty factor should be. Since we can minimize an SOS-convex polynomial by solving an SDP, Ahmadi et al. (2024) select the smallest penalty factor such that the regularized Taylor approximation is SOS-convex. That is, we solve the optimization problem

$$\begin{array}{ll} \min\limits_{t \in \mathbb{R}} & t \\ \text{s. t.} & x \mapsto T_{x_k,d}(x) + t \left\| x - x_k \right\|^{d'} \text{ is SOS-convex} \\ & t \geq 0. \end{array}$$

Note that this problem can be reformulated as an SDP.

7.2 Existence of an SOS-convex regularized Taylor approximation

To determine whether there always exists a $t \geq 0$ such that the polynomial

$$x \mapsto T_{x_k,d}(x) + t \left\| x - x_k \right\|^{d'}$$

is SOS-convex, we need to introduce a few lemmas.

Lemma 7.1 (Ahmadi et al., 2024)**.** Define the n-variate polynomial p of an even degree d for all $x \in \mathbb{R}^n$ by

$$p(x) = 1 + \left(\frac{d}{2} + 1 \right) \|x\|^d .$$

Then, there exists a $Q \succ 0$ such that, for all $x \in \mathbb{R}^n$, we have

$$p(x) = [x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}} .$$

Note that p is SOS. Therefore, we already know that there exists a $Q \succeq 0$ such that $p(x) = [x]_{\frac{d}{2}}^\top Q [x]_{\frac{d}{2}}$ for all $x \in \mathbb{R}^n$. However, this lemma states that we can take $Q \succ 0$.

Proof. We prove the lemma by induction on d.

If $d = 0$, then, for all $x \in \mathbb{R}^n$, we have

$$p(x) = 1 + 1 \cdot \|x\|^0 = 1 + 1 \cdot 1 = 2 = 1 \cdot 2 \cdot 1 = [x]_0 \cdot 2 \cdot [x]_0 ,$$

and $2 > 0$, so the lemma is true for $d = 0$.

Now, assume that the lemma is true for some even degree d. Then, there exists a $Q \succ 0$ such that, for all $x \in \mathbb{R}^n$, we have

$$1 + \left(\frac{d}{2} + 1\right) \|x\|^d = [x]_{\frac{d}{2}}^\top Q \, [x]_{\frac{d}{2}} = [x]_{\frac{d}{2}+1}^\top \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} [x]_{\frac{d}{2}+1} .$$

Let $P = \operatorname{diag}\left(\left(\binom{\frac{d}{2}+1}{\beta} \right)_{\beta \in \mathbb{N}_0^n, |\beta| = \frac{d}{2}+1} \right) \succ 0$. From the multinomial theorem, it then follows that, for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we have

$$\|x\|^{d+2} = \left(x_1^2 + \cdots + x_n^2\right)^{\frac{d}{2}+1} = \sum_{\substack{\beta \in \mathbb{N}_0^n \\ |\beta| = \frac{d}{2}+1}} \binom{\frac{d}{2}+1}{\beta} x^{2\beta} = [x]_{\frac{d}{2}+1}^\top \begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} [x]_{\frac{d}{2}+1} .$$

Let $A \in \mathbb{R}^{\binom{n+\frac{d}{2}}{\frac{d}{2}} \times \binom{n+\frac{d}{2}}{\frac{d}{2}+1}}$ such that, for all $x \in \mathbb{R}^n$, we have

$$-\left(\frac{d}{2} + 1\right) \|x\|^d = [x]_{\frac{d}{2}+1}^\top \begin{bmatrix} 0 & A \\ A^\top & 0 \end{bmatrix} [x]_{\frac{d}{2}+1} .$$

For all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$, we then have

$$1 + t \|x\|^{d+2} = [x]_{\frac{d}{2}+1}^\top \begin{bmatrix} Q & A \\ A^\top & tP \end{bmatrix} [x]_{\frac{d}{2}+1} .$$

Since $Q \succ 0$, it follows from Theorem A.19 that, for all $t \in \mathbb{R}$, we have $\begin{bmatrix} Q & A \\ A^\top & tP \end{bmatrix} \succ 0$ if and only if $tP \succ A^\top Q^{-1} A$. Since $P \succ 0$, there exists a $t > \frac{d}{2} + 2$ such that $tP \succ A^\top Q^{-1} A$. Then, $0 < \frac{\frac{d}{2}+2}{t} < 1$. For all $x \in \mathbb{R}^n$, we now have

$$\begin{aligned} 1 + \left(\frac{d}{2} + 2\right) \|x\|^{d+2} &= \frac{\frac{d}{2}+2}{t} \left(1 + t \|x\|^{d+2}\right) + \left(1 - \frac{\frac{d}{2}+2}{t}\right) \\ &= [x]_{\frac{d}{2}+1}^\top \left(\frac{\frac{d}{2}+2}{t} \begin{bmatrix} Q & A \\ A^\top & tP \end{bmatrix} + \begin{bmatrix} 1 - \frac{\frac{d}{2}+2}{t} & 0^\top \\ 0 & 0 \end{bmatrix} \right) [x]_{\frac{d}{2}+1} , \end{aligned}$$

and

$$\frac{\frac{d}{2}+2}{t} \begin{bmatrix} Q & A \\ A^\top & tP \end{bmatrix} + \begin{bmatrix} 1 - \frac{\frac{d}{2}+2}{t} & 0^\top \\ 0 & 0 \end{bmatrix} \succeq 0,$$

so the lemma is also true for $d + 2$.

Hence, the lemma is true for every even degree d. □

We can now prove the following lemma.

Lemma 7.2 (Ahmadi et al., 2024)**.** Let V be the vector space of n-variate polynomials of an even degree $d \geq 2$, and $\tilde{\mathcal{K}}$ the cone of SOS-convex polynomials in V. Define $p \in V$ for all $x \in \mathbb{R}^n$ by

$$p(x) = \|x\|^2 + \|x\|^d .$$

Then, $p \in \operatorname{int}\left(\tilde{\mathcal{K}}\right)$.

Proof. For all $x \in \mathbb{R}^n$, we have

$$\nabla p(x) = 2x + d\|x\|^{d-2} x,$$
$$\nabla^2 p(x) = 2I_n + d\|x\|^{d-2} I_n + d(d-2)\|x\|^{d-4} xx^\top.$$

Therefore, for all $x, y \in \mathbb{R}^n$, we have

$$y^\top \nabla^2 p(x)\, y = \left(1 + \frac{d}{2}\|x\|^{d-2}\right)\left(2\|y\|^2\right) + d(d-2)\|x\|^{d-4}\left(x^\top y\right)^2.$$

From Lemma 7.1, it follows that there exists a $Q \succ 0$ such that, for all $x \in \mathbb{R}^n$, we have

$$1 + \frac{d}{2}\|x\|^{d-2} = [x]_{\frac{d}{2}-1}^\top Q\, [x]_{\frac{d}{2}-1}.$$

From Theorem A.16 and Corollary A.17.1, it now follows that, for all $x, y \in \mathbb{R}^n$, we have

$$\begin{aligned}\left(1 + \frac{d}{2}\|x\|^{d-2}\right)\left(2\|y\|^2\right) &= \left([x]_{\frac{d}{2}-1}^\top Q\, [x]_{\frac{d}{2}-1}\right)\left(y^\top (2I_n)\, y\right) \\ &= \left([x]_{\frac{d}{2}-1} \otimes y\right)^\top (Q \otimes 2I_n)\left([x]_{\frac{d}{2}-1} \otimes y\right),\end{aligned}$$

and $Q \otimes 2I_n \succ 0$. Since the polynomial

$$x, y \mapsto d(d-2)\|x\|^{d-4}\left(x^\top y\right)^2$$

is SOS, there exists a $Q' \succeq 0$ such that, for all $x, y \in \mathbb{R}^n$, we have

$$d(d-2)\|x\|^{d-4}\left(x^\top y\right)^2 = \left([x]_{\frac{d}{2}-1} \otimes y\right)^\top Q'\left([x]_{\frac{d}{2}-1} \otimes y\right).$$

For all $x, y \in \mathbb{R}^n$, we now have

$$y^\top \nabla^2 p(x)\, y = \left([x]_{\frac{d}{2}-1} \otimes y\right)^\top (Q \otimes 2I_n + Q')\left([x]_{\frac{d}{2}-1} \otimes y\right),$$

and $Q \otimes 2I_n + Q' \succ 0$. Hence, $p \in \operatorname{int}\left(\tilde{\mathcal{K}}\right)$. □

The following theorem now shows that, if $\nabla^2 f(x_k) \succ 0$, then there always exists a $t \geq 0$ such that the polynomial

$$x \mapsto T_{x_k,d}(x) + t\|x_k - x\|^{d'}$$

is SOS-convex.

Theorem 7.3 (Ahmadi et al., 2024)**.** Let $f : D \to \mathbb{R}$ be a $d \geq 2$ times continuously differentiable function on an open domain $D \subseteq \mathbb{R}^n$, and let $B \subseteq D$ be a compact set. Assume that $\nabla^2 f(x) \succ 0$ for all $x \in B$. Then, there exists a $t \geq 0$ such that, for all $x \in B$, the polynomial

$$y \mapsto T_{x,d}(y) + t\|y - x\|^{d'},$$

where d' is the smallest even number greater than d, is SOS-convex.

Proof. Since $\nabla^2 f(x) \succ 0$ for all $x \in B$, and B is a compact set, there exists a $\delta > 0$ such that $\nabla^2 f(x) \succeq \delta I_n$ for all $x \in B$. For all $x \in B$, $u \in \mathbb{R}^n$ and $s > 0$, we now have

$$\begin{aligned}
T_{x,d}(x+su) &= \sum_{\alpha \in \mathbb{N}_d^n} \frac{D^\alpha f(x)}{\alpha!} s^{|\alpha|} u^\alpha \\
&= f(x) + s(\nabla f(x))^\top u + \frac{1}{2} s^2 u^\top \nabla^2 f(x) u + \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ |\alpha| \geq 3}} \frac{D^\alpha f(x)}{\alpha!} s^{|\alpha|} u^\alpha \\
&= f(x) + s(\nabla f(x))^\top u + \frac{1}{2} s^2 u^\top \left(\nabla^2 f(x) - \delta I_n\right) u \\
&\quad + \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ |\alpha| \geq 3}} \frac{D^\alpha f(x)}{\alpha!} s^{|\alpha|} u^\alpha + \frac{\delta}{2} s^2 \|u\|^2,
\end{aligned}$$

so

$$\begin{aligned}
\frac{2}{\delta}\frac{1}{s^2} T_{x,d}(x+su) + \|u\|^{d'} = &\frac{2}{\delta}\frac{1}{s^2} f(x) + \frac{2}{\delta}\frac{1}{s}(\nabla f(x))^\top u + \frac{1}{\delta} u^\top \left(\nabla^2 f(x) - \delta I_n\right) u \\
&+ \frac{2}{\delta} \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ |\alpha| \geq 3}} \frac{D^\alpha f(x)}{\alpha!} s^{|\alpha|-2} u^\alpha + \|u\|^2 + \|u\|^{d'}.
\end{aligned}$$

Let V' be the vector space of n-variate polynomials of degree d', and $\tilde{\mathcal{K}}'$ the cone of SOS-convex polynomials in V'. From Lemma 7.2, it then follows that $u \mapsto \|u\|^2 + \|u\|^{d'} \in \operatorname{int}\left(\tilde{\mathcal{K}}'\right)$. Therefore, there exists an $R > 0$ such that, for all $p \in V$ with $\|p\| \leq R$, we have $u \mapsto p(u) + \|u\|^2 + \|u\|^{d'} \in \tilde{\mathcal{K}}'$.

Since B is a compact set, there exists an $M > 0$ such that, for all $x \in B$, we have

$$\left\| u \mapsto \frac{2}{\delta} \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ |\alpha| \geq 3}} \frac{D^\alpha f(x)}{\alpha!} u^\alpha \right\| \leq M.$$

Let $s = \min\left\{1, \frac{R}{M}\right\}$, and $x \in B$. Then,

$$\left\| u \mapsto \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ |\alpha| \geq 3}} \frac{D^\alpha f(x)}{\alpha!} s^{|\alpha|-2} u^\alpha \right\| \leq s \left\| u \mapsto \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ |\alpha| \geq 3}} \frac{D^\alpha f(x)}{\alpha!} u^\alpha \right\| \leq sM \leq R.$$

Therefore, the polynomial

$$u \mapsto \sum_{\substack{\alpha \in \mathbb{N}_d^n \\ |\alpha| \geq 3}} \frac{D^\alpha f(x)}{\alpha!} s^{|\alpha|-2} u^\alpha + \|u\|^2 + \|u\|^{d'}$$

is SOS-convex. Since $\nabla^2 f(x) \succeq \delta I_n$, it follows from Theorem 4.4 that the quadratic function

$$u \mapsto \frac{2}{\delta}\frac{1}{s^2} f(x) + \frac{2}{\delta}\frac{1}{s}(\nabla f(x))^\top u + \frac{1}{\delta} u^\top \left(\nabla^2 f(x) - \delta I_n\right) u$$

is also SOS-convex. Hence, the polynomial

$$u \mapsto \frac{2}{\delta}\frac{1}{s^2}T_{x_k,d}\left(x+su\right)+\left\|u\right\|^{d'}$$

is SOS-convex. For all $y \in \mathbb{R}^n$ and $u = \frac{1}{s}\left(y-x\right)$, we have

$$\frac{\delta}{2}s^2\left(\frac{2}{\delta}\frac{1}{s^2}T_{x_k,d}\left(x+su\right)+\left\|u\right\|^{d'}\right) = T_{x_k,d}\left(y\right)+\frac{\delta}{2}\frac{1}{s^{d'-2}}\left\|y-x\right\|^{d'}.$$

Let $t = \frac{\delta}{2}\frac{1}{s^{d'-2}}$. Then, for all $x \in B$, the polynomial

$$y \mapsto T_{x,d}\left(y\right)+t\left\|y-x\right\|^{d'}$$

is SOS-convex. $\square$

If $\nabla^2 f\left(x_k\right) \not\succ 0$, then there might not exist a $t \geq 0$ such that the polynomial

$$x \mapsto T_{x_k,d}\left(x\right)+t\left\|x-x_k\right\|^{d'}$$

is SOS-convex (Ahmadi et al., 2024). Therefore, Ahmadi et al. (2024) instead consider the regularized Taylor approximation

$$x \mapsto T_{x_k,d}\left(x\right)+\frac{1}{2}\left(\delta-\lambda_{\min}\left(\nabla^2 f\left(x_k\right)\right)\right)^+\left\|x-x_k\right\|^2+t\left\|x-x_k\right\|^{d'},$$

where $\delta > 0$ is a small number, and solve the optimization problem

$$\begin{aligned} \min_{t\in\mathbb{R}} \quad & t \\ \text{s. t.} \quad & x \mapsto T_{x_k,d}\left(x\right)+\frac{1}{2}\left(\delta-\lambda_{\min}\left(\nabla^2 f\left(x_k\right)\right)\right)^+\left\|x-x_k\right\|^2+t\left\|x-x_k\right\|^{d'} \text{ is SOS-convex} \\ & t \geq 0. \end{aligned}$$

We now have

$$\nabla^2 f\left(x_k\right)+\left(\delta-\lambda_{\min}\left(\nabla^2 f\left(x_k\right)\right)\right)^+ I_n \succeq \delta I_n \succ 0,$$

so there always exists a $t \geq 0$ such that the polynomial

$$x \mapsto T_{x_k,d}\left(x\right)+\frac{1}{2}\left(\delta-\lambda_{\min}\left(\nabla^2 f\left(x_k\right)\right)\right)^+\left\|x-x_k\right\|^2+t\left\|x-x_k\right\|^{d'}$$

is SOS-convex. Define the function $t : D \to \mathbb{R}_+$ for all $x \in D$ by

$$t\left(x\right) = \min\left\{t \geq 0 : y \mapsto T_{x,d}\left(y\right)+\frac{1}{2}\left(\delta-\lambda_{\min}\left(\nabla^2 f\left(x\right)\right)\right)^+\left\|y-x\right\|^2+t\left\|y-x\right\|^{d'} \text{ is SOS-convex}\right\}.$$

Ahmadi et al. (2024) then define the function $\psi_{x_k,d} : \mathbb{R}^n \to \mathbb{R}$ for all $x \in \mathbb{R}^n$ by

$$\psi_{x_k,d}\left(x\right) = T_{x_k,d}\left(x\right)+\frac{1}{2}\left(\delta-\lambda_{\min}\left(\nabla^2 f\left(x_k\right)\right)\right)^+\left\|x-x_k\right\|^2+t\left(x_k\right)\left\|x-x_k\right\|^{d'}.$$

We now wish to set x_{k+1} equal to a minimizer of $\psi_{x_k,d}$.

7.3 Existence of a unique minimizer of the regularized Taylor approximation

A sufficient condition for a continuous function to have a minimizer is that the function is *coercive*.

Definition 7.1 (Coercive function)**.** A function $f : \mathbb{R}^n \to \mathbb{R}$, is called *coercive* if, for all $M \in \mathbb{R}$, there exists an $R > 0$ such that for all $x \in \mathbb{R}^n$ with $\|x\| > R$, we have $f(x) > M$, or

$$f(x) \to \infty \text{ as } \|x\| \to \infty.$$

Theorem 7.4 (e.g., Beck, 2014, Theorem 2.32)**.** Every continuous and coercive function has a minimizer.

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous and coercive function. Since f is coercive, there exists an $R > 0$ such that, for all $x \in \mathbb{R}^n$ with $\|x\| > R$, we have $f(x) > f(0)$. Since f is continuous, and the set $\{x \in \mathbb{R}^n : \|x\| \leq R\}$ is compact, it follows from Weierstrass' theorem that there exists an $x^* \in \mathbb{R}^n$ with $\|x^*\| \leq R$ such that

$$f(x^*) = \inf \{f(x) : x \in \mathbb{R}^n, \|x\| \leq R\}.$$

For all $x \in \mathbb{R}^n$ with $\|x\| > R$, we now have $f(x) > f(0) \geq f(x^*)$. Hence, x^* is a minimizer of f. $\square$

To prove that $\psi_{x_k,d}$ is coercive, Ahmadi et al. (2024) use the following quadrature rule for integration.

Theorem 7.5 (Clenshaw and Curtis, 1960)**.** Let p be a univariate polynomial of an even degree $d \geq 2$. Then, there exist $w_0, \ldots, w_d \geq 0$ and $s_0, \ldots, s_d \in [-1, 1]$, where $w_0 = \frac{1}{d^2-1}$ and $s_0 = -1$, such that

$$\int_{-1}^{1} p(s)\, ds = \sum_{i=0}^{d} w_i p(s_i).$$

The proof is beyond the scope of this thesis.

Ahmadi et al. (2024) now prove the following lemma.

Lemma 7.6 (Ahmadi et al., 2024)**.** Let $M : \mathbb{R} \to \mathbb{S}^{n\times n}$ be a function, where the entries of M are univariate polynomials of an even degree $d \geq 2$. Assume that $M(s) \succeq 0$ for all $s \in [0,1]$. Then,

$$\int_0^1 M(s)\, ds \succeq \frac{1}{2(d^2-1)} M(0).$$

Proof. From Theorem 7.5, it follows that there exist $w_0, \ldots, w_d \geq 0$ and $s_0, \ldots, s_d \in [-1,1]$, where $w_0 = \frac{1}{d^2-1}$ and $s_0 = -1$, such that

$$\begin{aligned}\int_0^1 M(s)\, ds &= \frac{1}{2}\int_{-1}^1 M\left(\frac{s+1}{2}\right) ds = \frac{1}{2}\sum_{i=0}^{d} w_i M\left(\frac{s_i+1}{2}\right) \\ &\succeq \frac{1}{2} w_0 M\left(\frac{s_0+1}{2}\right) = \frac{1}{2(d^2-1)} M(0).\end{aligned}$$

□

The following theorem now shows that $\psi_{x_k,d}$ has a unique minimizer.

Theorem 7.7 (Ahmadi et al., 2024)**.** Let p be a n-variate convex polynomial of an even degree $d \geq 4$. Assume that there exists an $x \in \mathbb{R}^n$ such that $\nabla^2 p(x) \succ 0$. Then, p has a unique minimizer.

Proof. Let $y \in \mathbb{R}^n$. Then (see also the proof of Theorem 4.8),

$$\begin{aligned} p(y) &= p(x) + (\nabla p(x))^\top (y-x) + (y-x)^\top \left(\int_0^1 \int_0^t \nabla^2 p(x + s(y-x))\, ds\, dt\right)(y-x) \\ &= p(x) + (\nabla p(x))^\top (y-x) + (y-x)^\top \left(\int_0^1 t \int_0^1 \nabla^2 p(x + ts(y-x))\, ds\, dt\right)(y-x).\end{aligned}$$

The entries of $\nabla^2 p$ are polynomials of degree $d-2 \geq 2$. Since p is convex, we have $\nabla^2 p(x + st(y-x)) \succeq 0$ for all $t, s \in [0,1]$. Therefore, from Lemma 7.6, it follows that, for all $t \in [0,1]$, we have

$$\int_0^1 \nabla^2 p(x + ts(y-x))\, ds \succeq \frac{1}{2\left((d-2)^2-1\right)} \nabla^2 p(x).$$

Hence,

$$\begin{aligned} p(y) &\geq p(x) + (\nabla p(x))^\top (y-x) + (y-x)^\top \left(\int_0^1 t \left(\frac{1}{2\left((d-2)^2-1\right)} \nabla^2 p(x)\right) dt\right)(y-x) \\ &= p(x) + (\nabla p(x))^\top (y-x) + (y-x)^\top \left(\frac{1}{4\left((d-2)^2-1\right)} \nabla^2 p(x)\right)(y-x).\end{aligned}$$

Since $\nabla^2 p(x) \succ 0$, we have $p(y) \to \infty$ as $\|y\| \to \infty$, so p is coercive. Since p is a polynomial, it is also continuous. Hence, p has a minimizer x^*.

Suppose that p has another minimizer $y^* \neq x^*$. Since p is a convex function, we then have $p(x^* + t(y^* - x^*)) = p(x^*)$ for all $t \in [0, 1]$. Since p is a polynomial, we now have $p(x^* + t(y^* - x^*)) = p(x^*)$ for all $t \in \mathbb{R}$. However, as $t \to \infty$, we now have $\|x^* + t(y^* - x^*)\| \to \infty$, but

$$p(x^* + t(y^* - x^*)) = p(x^*) \nrightarrow \infty,$$

so p is now not coercive. Hence, x^* is the only minimizer of p. □

Since $\nabla^2 \psi_{x_k,d}(x_k) = \nabla^2 f(x_k) + (\delta - \lambda_{\min}(\nabla^2 f(x_k)))^+ I_n \succ 0$, it follows directly that $\psi_{x_k,d}$ has a unique minimizer. Therefore, we can set x_{k+1} equal to the minimizer of $\psi_{x_k,d}$. We assume that $x_{k+1} \in D$.

Algorithm 7.1: d^{th}-order Newton method

Input: Objective function f, initial point x_0, small number $\delta > 0$, tolerance parameter $\varepsilon > 0$, maximum number of iterations N

for $k \leftarrow 0$ **to** $N - 1$ **do**
- **if** $\|\nabla f(x_k)\| < \varepsilon$ **then**
  - **return** x_k `// Converged`
- **function** $T_{x_k,d}(x)$**:**
  - `// Taylor approximation`
  - **return** $\sum_{\alpha \in \mathbb{N}^n_d} \frac{D^\alpha f(x_k)}{\alpha!} (x - x_k)^\alpha$
- $t(x_k) \leftarrow \min\left\{t \geq 0 : x \mapsto T_{x_k,d}(x) + \frac{1}{2}\left(\delta - \lambda_{\min}\left(\nabla^2 f(x_k)\right)\right)^+ \|x - x_k\|^2 + t\,\|x - x_k\|^{d'} \text{ is SOS-convex}\right\}$
- **function** $\psi_{x_k,d}(x)$**:**
  - `// Regularized Taylor approximation`
  - **return** $T_{x_k,d}(x) + \frac{1}{2}\left(\delta - \lambda_{\min}\left(\nabla^2 f(x_k)\right)\right)^+ \|x - x_k\|^2 + t(x_k)\,\|x - x_k\|^{d'}$
- $x_{k+1} \leftarrow \operatorname{argmin}\left\{\psi_{x_k,d}(x) : x \in \mathbb{R}^n\right\}$

return x_N `// Maximum number of iterations reached`

7.4 Convergence of the d^{th}-order Newton method

Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}$ is a d times continuously differentiable function on an open domain $D \subseteq \mathbb{R}^n$. We assume that the d^{th}-order partial derivatives of f are Lipschitz continuous.

To determine whether the d^{th}-order Newton method converges, we need the following lemma.

> **Lemma 7.8** (Ahmadi et al., 2024)**.** Let $f : D \to \mathbb{R}$ be a $d \geq 3$ times continuously differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that f has a local minimizer x^* with $\nabla^2 f(x^*) \succ 0$, and that, for all $\alpha \in \mathbb{N}_d^n$ with $|\alpha| = d$, the function $D^\alpha f$ is Lipschitz continuous with Lipschitz constant L. Then, there exists an $r > 0$ such that, for all $x \in D$ with $\|x - x^*\| \leq r$, we have
>
> $$\nabla^2 \psi_{x,d}(x^*) \succeq \frac{1}{2} \nabla^2 f(x^*).$$

Proof. From Theorem 3.8, it follows that, for all $i, j \in \{1, \ldots, n\}$ and $x \in D$, we have

$$|D_{ij} f(x^*) - D_{ij} T_{x,d}(x^*)| = |D_{ij} R_{x,d}(x^*)| \leq \frac{\sqrt{n^{d-2}} L}{(d-1)!} \|x - x^*\|^{d-1}.$$

Therefore,

$$\left\|\nabla^2 f(x^*) - \nabla^2 T_{x,d}(x^*)\right\| \to 0 \text{ as } x \to x^*.$$

Hence, there exists an $r > 0$ such that, for all $x \in D$ with $\|x - x^*\| \leq r$, we have

$$\left\|\nabla^2 f(x^*) - \nabla^2 T_{x,d}(x^*)\right\| \leq \frac{1}{2} \lambda_{\min}\left(\nabla^2 f(x^*)\right).$$

Let $x \in D$ with $\|x - x^*\| \leq r$. Then,

$$\nabla^2 f(x^*) - \nabla^2 T_{x,d}(x^*) \preceq \frac{1}{2} \nabla^2 f(x^*),$$

so $\nabla^2 T_{x,d}(x^*) \succeq \frac{1}{2} \nabla^2 f(x^*)$. Since, for all $y \in \mathbb{R}^n$, we have

$$\psi_{x,d}(y) = T_{x,d}(y) + t(x) \|y - x\|^{d'},$$

and the function $y \mapsto t(x) \|y - x\|^{d'}$ is convex, we have $\nabla^2 \psi_{x,d}(y) \succeq \nabla^2 T_{x,d}(y)$ for all $y \in \mathbb{R}^n$. Hence,

$$\nabla^2 \psi_{x,d}(x^*) \succeq \nabla^2 T_{x,d}(x^*) \succeq \frac{1}{2} \nabla^2 f(x^*).$$

□

It now follows that the d^{th}-order Newton method converges with order d to a local minimizer x^* of f if the initial point is sufficiently close to x^*

Theorem 7.9 (Local convergence with order d of the d^{th}-order Newton method; Ahmadi et al., 2024)**.** Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}$ is a d times continuously differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that f has a local minimizer x^* with $\nabla^2 f(x^*) \succ 0$, and that, for all $\alpha \in \mathbb{N}_d^n$ with $|\alpha| = d$, the function $D^\alpha f$ is Lipschitz continuous with Lipschitz constant L. Then, there exist an $r > 0$ and a $c > 0$ such that, for every sequence $(x_k)_{k=0}^\infty$ obtained by the d^{th}-order Newton method with $\|x_0 - x^*\| \leq r$, we have

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^d$$

for all $k \in \mathbb{N}_0$.

Proof. Since D is an open set, and $\nabla^2 f(x^*) \succ 0$, there exists an $r_1 > 0$ such that for all $x \in \mathbb{R}^n$, we have $x \in D$, and $\nabla^2 f(x) \succ 0$. From Lemma 7.8, it follows that there exists an $r_2 > 0$ such that, for all $x \in D$ with $\|x - x^*\| \leq r_2$, we have $\nabla^2 \psi_{x,d}(x^*) \succeq \frac{1}{2} \nabla^2 f(x^*)$.

Let $r' = \min\{r_1, r_2\}$. Since the set $\{x \in \mathbb{R}^n : \|x - x^*\| \leq r'\} \subseteq D$ is compact, it follows from Theorem 7.3 that there exists a t^* such that $t(x) \leq t^*$ for all $x \in \mathbb{R}^n$ with $\|x - x^*\| \leq r'$.

Let $(x_k)_{k=0}^\infty$ be a sequence obtained by the d^{th}-order Newton method, and let $k \in \mathbb{N}_0$. Assume that $\|x_k - x^*\| \leq r'$. From the chain rule and the fundamental theorem of calculus, it follows that

$$\nabla \psi_{x_k,d}(x_{k+1}) - \nabla \psi_{x_k,d}(x^*) = \left(\int_0^1 \nabla^2 \psi_{x_k,d}(x^* + t(x_{k+1} - x^*))\, dt \right) (x_{k+1} - x^*).$$

Since x_{k+1} is a minimizer of $\psi_{x_k,d}$, we have $\nabla \psi_{x_k,d}(x_{k+1}) = 0$. Therefore,

$$\nabla \psi_{x_k,d}(x^*) = -\left(\int_0^1 \nabla^2 \psi_{x_k,d}(x^* + t(x_{k+1} - x^*))\, dt \right) (x_{k+1} - x^*).$$

Since $\psi_{x_k,d}$ is convex, we have $\nabla^2 \psi_{x_k,d}(x^* + t(x_{k+1} - x^*))$ for all $t \in [0,1]$. From Lemma 7.6, it now follows that

$$\int_0^1 \nabla^2 \psi_{x_k,d}(x^* + t(x_{k+1} - x^*))\, dt \succeq \frac{1}{2\left((d' - 2)^2 - 1\right)} \nabla^2 \psi_{x_k,d}(x^*).$$

Therefore,

$$\begin{aligned}\left\|\nabla\psi_{x_k,d}\left(x^*\right)\right\| &\geq \lambda_{\min}\left(\int_0^1 \nabla^2\psi_{x_k,d}\left(x^*+t\left(x_{k+1}-x^*\right)\right)dt\right)\left\|x_{k+1}-x^*\right\| \\ &\geq \frac{1}{2\left(\left(d'-2\right)^2-1\right)}\lambda_{\min}\left(\nabla^2\psi_{x_k,d}\left(x^*\right)\right)\left\|x_{k+1}-x^*\right\| \\ &\geq \frac{1}{4\left(\left(d'-2\right)^2-1\right)}\lambda_{\min}\left(\nabla^2 f\left(x^*\right)\right)\left\|x_{k+1}-x^*\right\|.\end{aligned}$$

Hence,

$$\left\|x_{k+1}-x^*\right\| \leq \frac{4\left(\left(d'-2\right)^2-1\right)}{\lambda_{\min}\left(\nabla^2 f\left(x^*\right)\right)}\left\|\nabla\psi_{x_k,d}\left(x^*\right)\right\|.$$

For all $x\in\mathbb{R}^n$, we have

$$\begin{aligned}\psi_{x_k,d}\left(x\right) &= T_{x_k,d}\left(x\right)+t\left(x_k\right)\left\|x-x_k\right\|^{d'}, \\ \nabla\psi_{x_k,d}\left(x\right) &= \nabla T_{x_k,d}\left(x\right)+t\left(x_k\right)d'\left\|x-x_k\right\|^{d'-2}\left(x-x_k\right).\end{aligned}$$

Therefore,

$$\left\|\nabla\psi_{x_k,d}\left(x^*\right)\right\| \leq \left\|\nabla T_{x_k,d}\left(x^*\right)\right\|+t\left(x_k\right)d'\left\|x_k-x^*\right\|^{d'-1}.$$

Since x^* is a local minimizer of f, we have $\nabla f\left(x^*\right)=0$. From Theorem 3.8, it now follows that, for all $i\in\{1,\ldots,n\}$, we have

$$\left|D_iT_{x_k,d}\left(x^*\right)\right| = \left|D_iR_{x_k,d}\left(x^*\right)\right| \leq \frac{\sqrt{n^{d-1}}L}{d!}\left\|x_k-x^*\right\|^d.$$

Therefore,

$$\left\|\nabla T_{x_k,d}\left(x^*\right)\right\| \leq \frac{\sqrt{n^d}L}{d!}\left\|x_k-x^*\right\|^d.$$

Since $\left\|x_k-x^*\right\|\leq r'$, we have $t\left(x_k\right)\leq t^*$, and

$$\left\|x_k-x^*\right\|^{d'-1} = \begin{cases}\left\|x_k-x^*\right\|^d & \text{if } d \text{ is odd} \\ \left\|x_k-x^*\right\|^{d+1}\leq r'\left\|x_k-x^*\right\|^d & \text{if } d \text{ is even.}\end{cases}$$

Therefore,

$$\begin{aligned}\left\|\psi_{x_k,d}\left(x^*\right)\right\| &\leq \frac{\sqrt{n^d}L}{d!}\left\|x_k-x^*\right\|^d+t^*d'\max\left\{r',1\right\}\left\|x_k-x^*\right\|^d \\ &= \left(\frac{\sqrt{n^d}L}{d!}+t^*d'\max\left\{r',1\right\}\right)\left\|x_k-x^*\right\|^d\end{aligned}$$

Hence,

$$\left\|x_{k+1}-x^*\right\| \leq \frac{\lambda_{\min}\left(\nabla^2 f\left(x^*\right)\right)}{4\left(\left(d'-2\right)^2-1\right)}\left(\frac{\sqrt{n^d}L}{d!}+t^*d'\max\left\{r',1\right\}\right)\left\|x_k-x^*\right\|^d.$$

Let

$$c = \frac{\lambda_{\min}\left(\nabla^2 f\left(x^*\right)\right)}{4\left(\left(d'-2\right)^2-1\right)}\left(\frac{\sqrt{n^d}L}{d!}+t^*d'\max\left\{r',1\right\}\right),$$

and $r = \min\left\{r', \left(\frac{1}{c}\right)^{\frac{1}{d-1}}\right\}$. Then, $r^{d-1} \leq \frac{1}{c}$, so $cr^d \leq r$.

Let $(x_k)_{k=0}^{\infty}$ be a sequence obtained by the d^{th}-order Newton method with $\|x_0 - x^*\| \leq r$. If $\|x_k - x^*\| \leq r$ for some $k \in \mathbb{N}_0$, then

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^d \leq cr^d \leq r.$$

Hence, $\|x_k - x^*\| \leq r \leq r'$, so $\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^d$, for all $k \in \mathbb{N}_0$. $\square$

Corollary 7.9.1. Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}$ is a d times continuously differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that f has a local minimizer x^* with $\nabla^2 f(x) \succ 0$, and that, for all $\alpha \in \mathbb{N}_d^n$ with $|\alpha| = d$, the function $D^\alpha f$ is Lipschitz continuous with Lipschitz constant L. Then, there exists an $R > 0$ such that, for every sequence $(x_k)_{k=0}^{\infty}$ obtained by the d^{th}-order Newton method with $\|x_0 - x^*\| \leq R$, we have

$$\|x_k - x^*\| \leq 2R \left(\frac{1}{2}\right)^{d^k}$$

for all $k \in \mathbb{N}_0$.

Proof. From local convergence with order d of the d^{th}-order Newton method, it follows that there exist an $r > 0$ and a $c > 0$ such that, for every sequence $(x_k)_{k=0}^{\infty}$ obtained by the d^{th}-order Newton method with $\|x_0 - x^*\| \leq r$, we have $\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^d$ for all $k \in \mathbb{N}_0$.

Let $R = \min\left\{r, \frac{1}{2}\left(\frac{1}{c}\right)^{\frac{1}{d-1}}\right\}$. Then, $(2R)^{d-1} \leq \frac{1}{c}$, so $c(2R)^d \leq 2R$.

Let $(x_k)_{k=0}^{\infty}$ be a sequence obtained by the d^{th}-order Newton method with $\|x_0 - x^*\| \leq R$. If $\|x_k - x^*\| \leq 2R\left(\frac{1}{2}\right)^{d^k}$ for some $k \in \mathbb{N}_0$, then

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^d \leq c\left(2R\left(\frac{1}{2}\right)^{d^k}\right)^d = c(2R)^d \left(\frac{1}{2}\right)^{d^{k+1}} = 2R\left(\frac{1}{2}\right)^{d^{k+1}}.$$

Hence, $\|x_k - x^*\| \leq 2R\left(\frac{1}{2}\right)^{d^k}$ for all $k \in \mathbb{N}_0$. $\square$

7.5 Univariate case with $d = 3$

Ahmadi et al. (2024) show that, under some assumptions, the third-order Newton method for an unconstrained univariate optimization problem has a closed-form solution for the iterates.

Theorem 7.10 (Ahmadi et al., 2024)**.** Consider an unconstrained optimization problem

$$\inf_{x \in I} \quad f(x),$$

where $f : I \to \mathbb{R}$ is a three times continuously differentiable function on an interval $I \subseteq \mathbb{R}$. Let $(x_k)_{k=0}^{\infty}$ be a sequence obtained by the third-order Newton method.

Let $k \in \mathbb{N}_0$. Assume that $f''(x_k) > 0$ and $f'''(x_k) \neq 0$. Then,

$$x_{k+1} = x_k - \frac{2f''(x_k)}{f'''(x_k)} - \sqrt[3]{\frac{f'(x_k) - \frac{2(f''(x_k))^2}{3f'''(x_k)}}{\frac{(f'''(x_k))^2}{12f''(x_k)}}}.$$

Proof. For all $x \in \mathbb{R}$ and $t \geq 0$, we have

$$\begin{aligned} T_{x_k,3}(x) + t(x - x_k)^4 = f(x_k) + f'(x_k)(x - x_k) + \frac{f''(x_k)}{2}(x - x_k)^2 \\ + \frac{f'''(x_k)}{6}(x - x_k)^3 + t(x - x_k)^4. \end{aligned}$$

Therefore, for all $x, y \in \mathbb{R}$ and $t \geq 0$, we have

$$\begin{aligned} \frac{\partial^2}{\partial x^2}\left(T_{x_k,3}(x) + t(x - x_k)^4\right) y^2 &= \left(f''(x_k) + f'''(x_k)(x - x_k) + 12t(x - x_k)^2\right) y^2. \\ &= \left(f''(x_k) - \frac{(f'''(x_k))^2}{48t} + 12t\left(\frac{f'''(x_k)}{24t} + x - x_k\right)^2\right) y^2. \end{aligned}$$

For every $t \geq 0$, the polynomial

$$x, y \mapsto \left(f''(x_k) - \frac{(f'''(x_k))^2}{48t} + 12t\left(\frac{f'''(x_k)}{24t} + x - x_k\right)^2\right) y^2$$

is SOS if and only if $f''(x_k) - \frac{(f'''(x_k))^2}{48t} \geq 0$, or $t \geq \frac{(f'''(x_k))^2}{48f''(x_k)}$. Therefore, for every $t \geq 0$, the polynomial

$$x \mapsto T_{x_k,3}(x) + t(x - x_k)^4$$

is SOS-convex if and only if $t \geq \frac{(f'''(x_k))^2}{48f''(x_k)}$. Hence,

$$t(x_k) = \frac{(f'''(x_k))^2}{48f''(x_k)}.$$

For all $x \in \mathbb{R}$, we now have

$$\begin{aligned}\psi_{x_k,3}(x) &= f(x_k) + f'(x_k)(x - x_k) + \frac{f''(x_k)}{2}(x - x_k)^2 \\ &\quad + \frac{f'''(x_k)}{6}(x - x_k)^3 + \frac{(f'''(x_k))^2}{48 f''(x_k)}(x - x_k)^4,\end{aligned}$$

so

$$\begin{aligned}\psi'_{x_k,3}(x) &= f'(x_k) + f''(x_k)(x - x_k) + \frac{f'''(x_k)}{2}(x - x_k)^2 + \frac{(f'''(x_k))^2}{12 f''(x_k)}(x - x_k)^3 \\ &= f'(x_k) - \frac{2(f''(x_k))^2}{3 f'''(x_k)} + \frac{(f'''(x_k))^2}{12 f''(x_k)}\left(\frac{2 f''(x_k)}{f'''(x_k)} + x - x_k\right)^3,\end{aligned}$$

so

$$\psi'_{x_k,3}(x) = 0 \iff x = x_k - \frac{2 f''(x_k)}{f'''(x_k)} - \sqrt[3]{\frac{f'(x_k) - \frac{2(f''(x_k))^2}{3 f'''(x_k)}}{\frac{(f'''(x_k))^2}{12 f''(x_k)}}}.$$

Hence, since $\psi_{x_k,3}$ is a convex function, it is minimized by setting

$$x_{k+1} = x_k - \frac{2 f''(x_k)}{f'''(x_k)} - \sqrt[3]{\frac{f'(x_k) - \frac{2(f''(x_k))^2}{3 f'''(x_k)}}{\frac{(f'''(x_k))^2}{12 f''(x_k)}}}.$$

□

Below, we provide a visual illustration of an iteration of the third-order Newton method for an unconstrained univariate optimization problem. At iteration $k+1$, we replace the function f by its third-order Taylor approximation $T_{x_k,3}$ at x_k, add a quartic regularization term to obtain the SOS-convex polynomial $\psi_{x_k,3}$ of degree four, and set x_{k+1} equal to the unique minimizer of $\psi_{x_k,3}$.

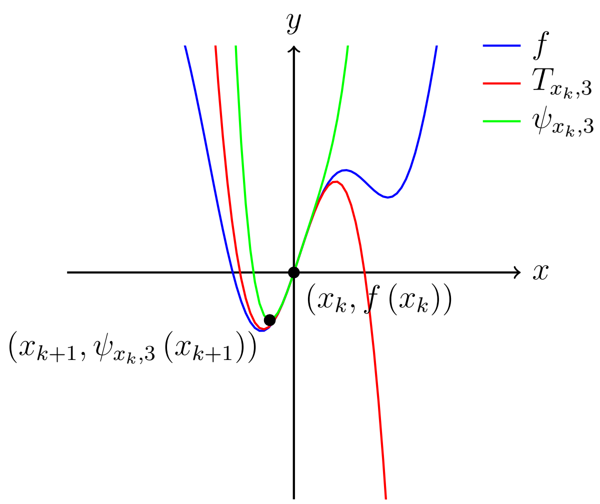

Figure 7.1: Illustration of the $(k+1)^{\text{th}}$ iteration of the third-order Newton method for an unconstrained univariate optimization problem.

7.6 Extension to constraints

Ahmadi et al. (2024) mention that the d^{th}-order Newton method can easily be extended to problems with SOS-convex polynomial constraints.

Consider a constrained optimization problem

$$\begin{aligned} \inf_{x \in D} \quad & f(x) \\ \text{s.t.} \quad & \sum_{\alpha \in \mathbb{N}_d^n} g_{i,\alpha} x^\alpha \leq 0, \quad i \in \{1, \ldots, m\} \end{aligned}$$

where f is a d times continuously differentiable functions on an open domain $D \subseteq \mathbb{R}^n$, and $g_1, \ldots, g_m$ are n-variate SOS-convex polynomials of degree d. We assume that

$$T = \left\{ x \in D : \sum_{\alpha \in \mathbb{N}_d^n} g_{1,\alpha} x^\alpha \leq 0, \ldots, \sum_{\alpha \in \mathbb{N}_d^n} g_{m,\alpha} x^\alpha \leq 0 \right\} \neq \emptyset.$$

A natural idea to update the current iterate x_k is to solve the optimization problem

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \quad & \psi_{x_k,d}(x) \\ \text{s.t.} \quad & \sum_{\alpha \in \mathbb{N}_d^n} g_{i,\alpha} x^\alpha \leq 0, \quad i \in \{1, \ldots, m\}, \end{aligned}$$

which is an SOS-convex polynomial optimization problem. To prove that this subproblem always has a unique minimizer, we need to extend Theorem 7.4 and Theorem 7.7 to constrained cases.

Theorem 7.11. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous and coercive function, and $T \subseteq \mathbb{R}^n$ a nonempty closed set. Then, there exists an $x^* \in T$ such that

$$f(x^*) = \inf \{f(x) : x \in T\}.$$

Proof. Let $x_0 \in T$. Since f is coercive, there exists an $R > \|x_0\|$ such that, for all $x \in \mathbb{R}^n$ with $\|x\| > R$, we have $f(x) > f(x_0)$. Since f is continuous, and the set $\{x \in T : \|x\| \leq R\}$ is compact, it follows from Weierstrass' theorem that there exists an $x^* \in T$ with $\|x^*\| \leq R$ such that

$$f(x^*) = \inf \{f(x) : x \in T, \|x\| \leq R\}.$$

For all $x \in T$ with $\|x\| > R$, we now have $f(x) > f(x_0) \geq f(x^*)$. Hence,

$$f(x^*) = \inf \{f(x) : x \in T\}.$$

□

Theorem 7.12. Let p be a n-variate convex polynomial of an even degree $d \geq 4$, and $T \subseteq \mathbb{R}^n$ a nonempty closed convex set. Assume that there exists an $x \in \mathbb{R}^n$ such that $\nabla^2 p(x) \succ 0$. Then, there exists a unique $x^* \in T$ such that

$$p(x^*) = \inf \{p(x) : x \in T\}.$$

Proof. From the proof of Theorem 7.7, it follows that p is continuous and coercive. From Theorem 7.11, it now follows that there exists an $x^* \in T$ such that

$$p(x^*) = \inf \{p(x) : x \in T\}.$$

Suppose that there exists a $y^* \in T$ with $y^* \neq x^*$ such that $p(y^*) = p(x^*)$. Since p is a convex function, and T a convex set, we then have $p(x^* + t(y^* - x^*)) = p(x^*)$ for all $t \in [0,1]$. However, from the proof of Theorem 7.7, it now follows that p is not coercive. Hence, x^* is unique. □

Since the feasible region

$$\left\{ x \in \mathbb{R}^n : \sum_{\alpha \in \mathbb{N}_d^n} g_{1,\alpha} x^\alpha \leq 0, \ldots, \sum_{\alpha \in \mathbb{N}_d^n} g_{m,\alpha} x^\alpha \leq 0 \right\} \supseteq T$$

is nonempty, closed and convex, the subproblem

$$\begin{array}{ll} \inf\limits_{x \in \mathbb{R}^n} & \psi_{x_k,d}(x) \\ \text{s.t.} & \sum\limits_{\alpha \in \mathbb{N}^n_{\frac{d}{2}}} g_{i,\alpha} x^\alpha \leq 0, \quad i \in \{1, \ldots, m\} \end{array}$$

always has a unique minimizer, and we can set x_{k+1} equal to this minimizer. We assume that $x_{k+1} \in D$.

Note that, for a minimizer x^* of a constrained optimization problem, we need not have $\nabla f(x^*) = 0$. Therefore, we need different stopping criterion than $\|\nabla f(x_k)\| < \varepsilon$ for some $\varepsilon > 0$.

For a constrained optimization problem, the *Karush-Kuhn-Tucker (KKT) conditions* are necessary conditions for a point x^* to be a local minimizer. For a convex optimization problem, the KKT conditions are also sufficient for x^* to be a global minimizer (Boyd and Vandenberghe, 2004).

Theorem 7.13 (Karush-Kuhn-Tucker conditions; e.g., Boyd and Vandenberghe, 2004, Subsection 5.5.3)**.** Consider a constrained optimization problem

$$\begin{aligned} \inf_{x \in D} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i \in \{1, \dots, m\} \end{aligned} \tag{P}$$

where $f, g_1, \dots, g_m : D \to \mathbb{R}$ are functions on an open domain $D \subseteq \mathbb{R}^n$. Assume that (P) has a local minimizer x^*. Then, there exist $\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}$ such that

$$\begin{cases} \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0 & \\ \lambda_i^* g_i(x^*) = 0, & i \in \{1, \dots, m\} \\ g_i(x^*) \leq 0, & i \in \{1, \dots, m\} \\ \lambda_i^* \geq 0, & i \in \{1, \dots, m\}. \end{cases}$$

Note that the KKT conditions form a system of linear equations and inequalities. Therefore, we can determine whether a point $x^* \in D$ satisfies the KKT conditions by solving an LP.

The constraints of our subproblem ensure that $\sum_{\alpha \in \mathbb{N}_d^n} g_{i,\alpha} x_k^\alpha \leq 0$ for all $i \in \{1, \dots, m\}$. We now terminate the algorithm when there exist $\lambda_1, \dots, \lambda_m \geq 0$ such that

$$\begin{cases} \left| D^{e_j} f(x_k) + \sum_{i=1}^{m} \lambda_i \left(\sum_{\alpha \in \mathbb{N}_{d-1}^n} \frac{(\alpha + e_j)!}{\alpha!} g_{i,\alpha+e_j} x_k^\alpha \right) \right| \leq \varepsilon, & j \in \{1, \dots, n\} \\ \left| \lambda_i \left(\sum_{\alpha \in \mathbb{N}_d^n} g_{i,\alpha} x_k^\alpha \right) \right| \leq \varepsilon, & i \in \{1, \dots, m\}. \end{cases}$$

Algorithm 7.2: d^{th}-order Newton method with SOS-convex polynomial constraints

Input: Objective function f, SOS-convex polynomial constraints $g_1, \ldots, g_m$ of degree d, initial point x_0, small number $\delta > 0$, tolerance parameter $\varepsilon > 0$, maximum number of iterations N

for $k \leftarrow 0$ **to** $N - 1$ **do**

`// KKT conditions`

if $\left\{\begin{array}{l}\lambda_1, \ldots, \lambda_m \geq 0 : \\ \left| D^{e_j} f(x_k) + \sum_{i=1}^{m} \lambda_i \left(\sum_{\alpha \in \mathbb{N}^n_{d-1}} \frac{(\alpha + e_j)!}{\alpha!} g_{i,\alpha+e_j} x_k^\alpha \right) \right| \leq \varepsilon, \ j \in \{1, \ldots, n\} \\ \left| \lambda_i \left(\sum_{\alpha \in \mathbb{N}^n_d} g_{i,\alpha} x_k^\alpha \right) \right| \leq \varepsilon, \ i \in \{1, \ldots, m\} \end{array}\right\} \neq \emptyset$ **then**

return x_k `// Converged`

function $T_{x_k,d}(x)$**:**

`// Taylor approximation`

return $\sum_{\alpha \in \mathbb{N}^n_d} \frac{D^\alpha f(x_k)}{\alpha!} (x - x_k)^\alpha$

$$t(x_k) \leftarrow \min \Big\{ t \geq 0 : x \mapsto T_{x_k,d}(x) + \tfrac{1}{2} \left(\delta - \lambda_{\min} \left(\nabla^2 f(x_k) \right) \right)^+ \|x - x_k\|^2 + t \|x - x_k\|^{d'} \text{ is SOS-convex} \Big\}$$

function $\psi_{x_k,d}(x)$**:**

`// Regularized Taylor approximation`

return $T_{x_k,d}(x) + \frac{1}{2} \left(\delta - \lambda_{\min} \left(\nabla^2 f(x_k) \right) \right)^+ \|x - x_k\|^2 + t(x_k) \|x - x_k\|^{d'}$

$$x_{k+1} \leftarrow \operatorname{argmin} \left\{ \psi_{x_k,d}(x) : x \in \mathbb{R}^n, \sum_{\alpha \in \mathbb{N}^n_{\frac{d}{2}}} g_{1,\alpha} x^\alpha \leq 0, \ldots, \sum_{\alpha \in \mathbb{N}^n_{\frac{d}{2}}} g_{m,\alpha} x^\alpha \leq 0 \right\}$$

return x_N `// Maximum number of iterations reached`

Chapter 8

Conclusion

The d^{th}-order Newton method from Ahmadi et al. (2024) is a variant of Newton's method that, at each iteration, minimizes an SOS-convex regularized d^{th}-order Taylor approximation. Ahmadi et al. (2024) have proven that such a regularized Taylor approximation always exists and has a unique minimizer. Moreover, both the optimal regularized Taylor approximation and its minimizer can be found by solving a semidefinite program.

The classical Newton method exhibits local quadratic convergence (i.e., convergence of order two) around a local minimizer, whereas the d^{th}-order Newton method exhibits local convergence of order d. Therefore, in terms of the number of iterations required for an optimization algorithm to converge, a higher-order Newton method is preferable to the classical method.

Furthermore, we successfully extended the d^{th}-order Newton method to optimization problems with SOS-convex polynomial constraints. At each iteration, we can minimize an SOS-convex regularized Taylor approximation subject to the problem's constraints. We have shown that the feasible region is nonempty and that each subproblem has a unique optimal solution.

8.1 Future research

In this thesis, we were able to extend the higher-order Newton method from Ahmadi et al. (2024) to problems with SOS-convex constraints. A natural idea for more general constraints would be to replace each constraint by either an SOS-convex regularized Taylor approximation or its affine approximation. However, ensuring the feasibility of iterates could pose a challenge.

For the unconstrained d^{th}-order Newton method, local convergence of order d has been established. Therefore, analyzing whether this result generalizes to the constrained case presents an interesting direction for future research.

Ahmadi et al. (2024) observed empirically that the basin of attraction around a local minimizer, i.e., the region in which the initial point must lie for the algorithm to converge, tends to be larger for a higher-order Newton method than for the classical method. However, to the best of our knowledge, this has not yet been theoretically proven.

Lastly, Ahmadi et al. (2024) proved that the higher-order Newton method exhibits local convergence for a general objective function with a local minimizer. Therefore, an interesting question is whether we can achieve global convergence for a (strongly) convex objective function.

Appendix A

Linear algebra

A.1 Inner product spaces and normed vector spaces

Let V be a vector space. We can then define an *inner product* on V.

Definition A.1 (Inner product)**.** An *inner product* on a vector space V is an operator $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ such that, for all $u, v, w \in V$ and $c \in \mathbb{R}$, we have

- symmetry: $\langle u, v \rangle = \langle v, u \rangle$,
- additivity: $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,
- homogeneity: $\langle cu, v \rangle = c \langle u, v \rangle$,
- positive definiteness: $\langle u, u \rangle \geq 0$, and $\langle u, u \rangle = 0$ if and only if $u = 0$.

A vector space on which an inner product is defined is called an *inner product space.*

In this thesis, we use the following inner products on $\mathbb{R}^n$ and $\mathbb{R}^{m \times n}$.

Definition A.2 (Inner product on $\mathbb{R}^n$)**.** The standard inner product on $\mathbb{R}^n$ is defined for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ by

$$\langle x, y \rangle = x^\top y = x_1 y_1 + \cdots + x_n y_n.$$

Definition A.3 (Inner product on $\mathbb{R}^{m\times n}$)**.** The standard inner product on $\mathbb{R}^{m\times n}$ is defined for all

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n} \text{ and } B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

by

$$\langle A, B\rangle = \operatorname{tr}\left(A^\top B\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}.$$

We can also define a *norm* on a vector space V.

Definition A.4 (Norm)**.** A *norm* on a vector space V is an operator $\|\cdot\| : V \to \mathbb{R}$ such that, for all $u, v \in V$ and $c \in \mathbb{R}$, we have

- positive definiteness: $\|u\| \geq 0$, and $\|u\| = 0$ if and only if $u = 0$,
- absolute homogeneity: $\|cu\| = |c|\,\|u\|$,
- triangle inequality: $\|u + v\| \leq \|u\| + \|v\|$.

A vector space on which a norm is defined is called a *normed vector space.*

We use the following norm on an inner product space.

Definition A.5 (Norm on an inner product space)**.** The standard norm on an inner product space V is defined for all $u \in V$ by

$$\|u\| = \sqrt{\langle u, u\rangle}.$$

Positive definiteness and absolute homogeneity of the standard norm on an inner product space follow directly from symmetry, homogeneity and positive definiteness of the inner product.

To prove the triangle inequality, we first need the *Cauchy-Schwarz inequality.*

Theorem A.1 (Cauchy-Schwarz inequality; e.g. Lay et al., 2021, Section 6.7, Theorem 16)**.** Let V be an inner product space. For all $u, v \in V$, we then have

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|\,.$$

Proof. Let $u, v \in V$. If $v = 0$, then $|\langle u, v\rangle| = \|u\|\,\|v\| = 0$.

Now, assume that $v \neq 0$. Then,

$$\begin{aligned}
\left\| u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v \right\|^2 &= \left\langle u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v, u - \frac{\langle u, v\rangle}{\langle v, v\rangle} v \right\rangle \\
&= \langle u, u\rangle - 2\frac{\langle u, v\rangle}{\langle v, v\rangle} \langle u, v\rangle + \frac{(\langle u, v\rangle)^2}{(\langle v, v\rangle)^2} \langle v, v\rangle \\
&= \langle u, u\rangle - 2\frac{(\langle u, v\rangle)^2}{\langle v, v\rangle} + \frac{(\langle u, v\rangle)^2}{\langle v, v\rangle} \\
&= \langle u, u\rangle - \frac{(\langle u, v\rangle)^2}{\langle v, v\rangle} = \|u\|^2 - \frac{(\langle u, v\rangle)^2}{\|v\|^2} \geq 0.
\end{aligned}$$

Therefore, $(\langle u, v\rangle)^2 \leq \|u\|^2 \|v\|^2$. Hence, $|\langle u, v\rangle| \leq \|u\| \|v\|$. □

We can now prove the triangle inequality.

Theorem A.2 (Triangle inequality; e.g. Lay et al., 2021, Section 6.7, Theorem 17)**.** Let V be an inner product space. For all $u, v \in V$, we then have

$$\|u + v\| \leq \|u\| + \|v\| .$$

Proof. Let $u, v \in V$. From the Cauchy-Schwarz inequality, it then follows that

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v\rangle = \langle u, u\rangle + 2\langle u, v\rangle + \langle v, v\rangle \\
&\leq \|u\|^2 + 2\|u\| \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 .
\end{aligned}$$

Hence, $\|u + v\| \leq \|u\| + \|v\|$. □

We use the following norms on $\mathbb{R}^n$ and $\mathbb{R}^{m \times n}$.

Definition A.6 (Norm on $\mathbb{R}^n$)**.** The *Euclidean* norm on $\mathbb{R}^n$ is defined for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ by

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Definition A.7 (Norm on $\mathbb{R}^{m \times n}$)**.** The *spectral* norm on $\mathbb{R}^{m \times n}$ is defined for all $A \in \mathbb{R}^{m \times n}$ by

$$\|A\| = \sigma_{\max}(A) = \sqrt{\lambda_{\max}(A^\top A)} = \max\{\|Au\| : u \in \mathbb{R}^n, \|u\| = 1\} .$$

It can be directly verified that that the spectral norm on $\mathbb{R}^{m \times n}$ satisfies positive definiteness and absolute homogeneity.

We now prove the triangle inequality.

Theorem A.3 (Triangle inequality)**.** For all $A, B \in \mathbb{R}^{m\times n}$, we have

$$\|A + B\| \leq \|A\| + \|B\| .$$

Proof. Let $A, B \in \mathbb{R}^{m\times n}$. From the triangle inequality on $\mathbb{R}^m$, it then follows that

$$\begin{aligned} \|A + B\| &= \max\left\{\|(A+B)\,u\| : u \in \mathbb{R}^n, \|u\| = 1\right\} \\ &\leq \max\left\{\|Au\| + \|Bu\| : u \in \mathbb{R}^n, \|u\| = 1\right\} \\ &\leq \max\left\{\|Au\| : u \in \mathbb{R}^n, \|u\| = 1\right\} + \max\left\{\|Bu\| : u \in \mathbb{R}^n, \|u\| = 1\right\} \\ &= \|A\| + \|B\| . \end{aligned}$$

□

The spectral norm on $\mathbb{R}^{m\times n}$ and the Euclidean norm on $\mathbb{R}^n$ are submultiplicative.

Theorem A.4 (Submultiplicativity)**.** For all $A \in \mathbb{R}^{m\times n}$ and $x \in \mathbb{R}^n$, we have

$$\|Ax\| \leq \|A\|\,\|x\| .$$

Proof. Let $A \in \mathbb{R}^{m\times n}$ and $x \in \mathbb{R}^n$. Then,

$$\|Ax\|^2 = x^\top A^\top Ax \leq \lambda_{\max}\left(A^\top A\right)\|x\|^2 = \|A\|^2\,\|x\|^2 .$$

Hence, $\|Ax\| \leq \|A\|\,\|x\|$. □

A.2 Orthogonal and orthonormal vectors

Two vectors in an inner product space are called *orthogonal* if their inner product equals zero.

Definition A.8 (Orthogonal vectors)**.** Let V be an inner product space, and $u, v \in V$. Then, u and v are called *orthogonal* if

$$\langle u, v\rangle = 0.$$

The following theorem states that nonzero orthogonal vectors are always linearly independent.

Theorem A.5 (e.g. Lay et al., 2021, Section 6.2, Theorem 4)**.** Let V be an inner product space, and $u_1, \dots, u_n \in V$. Assume that $u_1, \dots, u_n$ are nonzero and orthogonal. Then, $u_1, \dots, u_n$ are linearly independent.

Proof. Since $u_1, \dots, u_n$ are orthogonal, we have $\langle u_i, u_j \rangle = 0$ for all $i, j \in \{1, \dots, n\}$ with $i \neq j$.

Let $c_1, \dots, c_n \in \mathbb{R}$ such that $c_1 u_1 + \cdots + c_n u_n = 0$. For all $i \in \{1, \dots, n\}$, we then have

$$\langle c_1 u_1 + \cdots + c_n u_n, u_i \rangle = c_1 \langle u_1, u_i \rangle + \cdots + c_n \langle u_n, u_i \rangle = c_i \langle u_i, u_i \rangle = 0,$$

so $c_i = 0$. Hence, $u_1, \dots, u_n$ are linearly independent. □

Corollary A.5.1. Let V be an n-dimensional inner product space, and $u_1, \dots, u_n \in V$. Assume that $u_1, \dots, u_n$ are nonzero and orthogonal. Then, the set $\{u_1, \dots, u_n\}$ is a basis of V.

The following theorem states that every finite-dimensional inner product space has an orthogonal basis.

Theorem A.6 (Gram–Schmidt theorem, e.g. Lay et al., 2021, Section 6.4, Theorem 11)**.** Let $\{u_1, \dots, u_n\}$ be a basis of an inner product space V. For every $i \in \{1, \dots, n\}$, let

$$v_i = u_i - \frac{\langle u_i, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \cdots - \frac{\langle u_i, v_{i-1} \rangle}{\langle v_{i-1}, v_{i-1} \rangle} v_{i-1}.$$

Then, the set $\{v_1, \dots, v_n\}$ is an orthogonal basis of V.

Proof. We prove by induction that, for all $i \in \{1, \dots, n\}$, the set $\{v_1, \dots, v_i\}$ is an orthogonal basis of $\operatorname{Span}\{u_1, \dots, u_i\}$. We have $v_1 = u_1 \neq 0$, so the set $\{v_1\}$ is an orthogonal basis of $\operatorname{Span}\{u_1\}$.

Assume that, the set $\{v_1, \dots, v_i\}$ is an orthogonal basis of $\operatorname{Span}\{u_1, \dots, u_i\}$. For all $j \in \{1, \dots, i\}$, we then have

$$\begin{aligned}
\langle v_{i+1}, v_j \rangle &= \left\langle u_{i+1} - \frac{\langle u_{i+1}, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \cdots - \frac{\langle u_{i+1}, v_i \rangle}{\langle v_i, v_i \rangle} v_i, v_j \right\rangle \\
&= \langle u_{i+1}, v_j \rangle - \frac{\langle u_{i+1}, v_1 \rangle}{\langle v_1, v_1 \rangle} \langle v_1, v_j \rangle - \cdots - \frac{\langle u_{i+1}, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_j \rangle \\
&= \langle u_{i+1}, v_j \rangle - \frac{\langle u_{i+1}, v_j \rangle}{\langle v_j, v_j \rangle} \langle v_j, v_j \rangle = \langle u_{i+1}, v_j \rangle - \langle u_{i+1}, v_j \rangle = 0.
\end{aligned}$$

Hence, $v_1, \dots, v_i, v_{i+1}$ are orthogonal. Since $u_1, \dots, u_i, u_{i+1}$ are linearly independent, we have

$$u_{i+1} \notin \operatorname{Span}\{u_1, \dots, u_i\} = \operatorname{Span}\{v_1, \dots, v_i\}.$$

Therefore, $v_{i+1} \neq 0$, and

$$v_1, \dots, v_i, v_{i+1} \in \operatorname{Span}\{v_1, \dots, v_i, u_{i+1}\} = \operatorname{Span}\{u_1, \dots, u_i, u_{i+1}\}.$$

Hence, $\{v_1, \dots, v_i, v_{i+1}\}$ is an orthogonal basis of $\operatorname{Span}\{u_1, \dots, u_i, u_{i+1}\}$. □

Vectors in an inner product space are called *orthonormal* if they are orthogonal, and their norms equal one.

Definition A.9 (Orthonormal vectors)**.** Let V be an inner product space, and $u_1, \dots, u_n \in V$. Then, $u_1, \dots, u_n$ are called *orthonormal* if they are orthogonal, and

$$\|u_1\| = \dots = \|u_n\| = 1.$$

Since every finite-dimensional inner product space has an orthogonal basis, it also has an orthonormal basis.

Theorem A.7. Let $\{u_1, \dots, u_n\}$ be an orthogonal basis of an inner product space V. Then, the set $\left\{\frac{u_1}{\|u_1\|}, \dots, \frac{u_n}{\|u_n\|}\right\}$ is an orthonormal basis of V.

The following theorem now states that we can express the norm of a vector in an inner product space in terms of its coordinates with respect to an orthonormal basis.

Theorem A.8. Let $\{u_1, \dots, u_n\}$ be an orthonormal basis of an inner product space V. Let $v \in V$, and $c_1, \dots, c_n \in \mathbb{R}$ such that $v = c_1 u_1 + \dots + c_n u_n$. Then,

$$\|v\| = \sqrt{c_1^2 + \dots + c_n^2}.$$

Proof. Since $u_1, \dots, u_n$ are orthonormal, it follows that,

$$\begin{aligned}
\|v\|^2 &= \left\|\sum_{i=1}^n c_i u_i\right\|^2 = \left\langle \sum_{i=1}^n c_i u_i, \sum_{j=1}^n c_j u_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n c_i c_j \langle u_i, u_j \rangle \\
&= \sum_{i=1}^n c_i \langle u_i, u_i \rangle = \sum_{i=1}^n c_i \|u_i\|^2 = \sum_{i=1}^n c_i.
\end{aligned}$$

Hence, $\|v\| = \sqrt{c_1^2 + \dots + c_n^2}$. $\square$

Corollary A.8.1. Let $\{u_1, \dots, u_n\}$ be an orthonormal basis of an inner product space V. Let $v \in V$, and $c_1, \dots, c_n \in \mathbb{R}$ such that $v = c_1 u_1 + \dots + c_n u_n$. Then,

$$|c_1|, \dots, |c_n| \leq \|v\|.$$

Proof. Let $i \in \{1, \dots, n\}$. Then, $c_i^2 \leq c_1^2 + \dots + c_n^2 = \|v\|^2$, so $|c_i| \leq \|x\|$. $\square$

A.3 Convex cones

A set $\mathcal{K}$ in a vector space V is called a *cone* if $\mathcal{K}$ is closed under nonnegative scalar multiplication.

Definition A.10 (Cone)**.** Let V be a vector space, and $\mathcal{K} \subseteq V$. Then, $\mathcal{K}$ is called a *cone* if, for all $x \in \mathcal{K}$ and $\lambda \geq 0$, we have

$$\lambda x \in \mathcal{K}.$$

It follows directly that $\mathcal{K}$ is a convex cone if and only if $\mathcal{K}$ is closed under nonnegative linear combinations.

Theorem A.9. Let V be a vector space, and $\mathcal{K} \subseteq V$. Then, $\mathcal{K}$ is a convex cone if and only if, for all $x, y \in \mathcal{K}$ and $\lambda, \mu \geq 0$, we have

$$\lambda x + \mu y \in \mathcal{K}.$$

The linear image of a convex cone is also a convex cone.

Theorem A.10. Let V and W be vector spaces, $T : V \to W$ a linear transformation, and $\mathcal{K}$ a convex cone in V. Then, $T(\mathcal{K})$ is also a convex cone.

Proof. Let $u, v \in T(\mathcal{K})$ and $\lambda, \mu \geq 0$. Then, there exist $x, y \in \mathcal{K}$ such that $u = T(x)$ and $y = T(y)$. Since $\mathcal{K}$ is a convex cone, we have $\lambda x + \mu y \in \mathcal{K}$. Since T is a linear transformation, it now follows that

$$\lambda u + \mu v = \lambda T(x) + \mu T(y) = T(\lambda x + \mu y) \in T(\mathcal{K}).$$

Hence, $T(\mathcal{K})$ is also a convex cone. □

The reverse is also true.

Theorem A.11. Let V and W be vector spaces, $T : V \to W$ a linear transformation, and $\mathcal{L}$ a convex cone in W. Then, the set

$$\mathcal{K} = \{u \in V : T(u) \in \mathcal{L}\}$$

is also a convex cone.

Proof. Let $u, v \in \mathcal{K}$ and $\lambda, \mu \geq 0$. Then, $T(u) \in \mathcal{L}$ and $T(v) \in \mathcal{L}$. Since T is a linear transformation, and $\mathcal{L}$ a convex cone, we now have

$$T(\lambda u + \mu v) = \lambda T(u) + \mu T(v) \in \mathcal{L},$$

so $\lambda u + \mu v \in \mathcal{K}$. Hence, $\mathcal{K}$ is also a convex cone. □

Let V be an inner product space, and $K \subseteq V$. Then, the set of points that have a nonnegative inner product with all points in $\mathcal{K}$ is a convex cone, and is called the *dual cone* of $\mathcal{K}$. Note that $\mathcal{K}$ itself need not be a cone.

Definition A.11 (Dual cone)**.** Let V be an inner product space, and $K \subseteq V$. Then, the set

$$K^* = \{s \in V : \langle s, x \rangle \geq 0 \text{ for all } x \in \mathcal{K}\}$$

is called the *dual cone* of $\mathcal{K}$.

Theorem A.12. The dual cone of a set in an inner product space is a convex cone.

Proof. Let V be an inner product space, $K \subseteq V$, and $\mathcal{K}^*$ the dual cone of $\mathcal{K}$. Let $s, t \in \mathcal{K}^*$ and $\lambda, \mu \geq 0$. Since $s, t \in \mathcal{K}^*$, it follows that $\langle s, x \rangle \geq 0$ and $\langle t, x \rangle \geq 0$ for all $x \in \mathcal{K}$. Therefore,

$$\langle \lambda s + \mu t, x \rangle = \lambda \langle s, x \rangle + \mu \langle t, x \rangle \geq 0$$

for all $x \in \mathcal{K}$, so $\lambda s + \mu t \in \mathcal{K}^*$. Hence, $\mathcal{K}^*$ is a convex cone. $\square$

A.4 Conic programming

An optimization problem with a linear objective function and linear constraints, where we optimize over points in a cone, is called a *conic program.*

Definition A.12 (Conic program)**.** A *conic program* is an optimization problem

$$\begin{array}{ll} \inf\limits_{X \in V} & \langle C, X \rangle \\ \text{s.t.} & \langle A_i, X \rangle \leq b_i, \quad i \in \{1, \dots, m\} \\ & X \in \mathcal{K}, \end{array} \qquad \text{(P)}$$

where $\mathcal{K}$ is a cone in an inner product space V, and where $A_1, \dots, A_m \in V$, $b_1, \dots, b_m \in \mathbb{R}$ and $C \in V$.

If $\mathcal{K} = \mathbb{R}_+^n$, then (P) is called a *linear program*.

Definition A.13. A *linear program (LP)* is an optimization problem

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \quad & c^\top x \\ \text{s.t.} \quad & Ax \le b \\ & x \ge 0, \end{aligned} \tag{LP}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}^n$.

If $\mathcal{K} = \mathbb{S}_+^{n \times n}$, then (P) is called a *semidefinite program*.

Definition A.14. A *semidefinite program (SDP)* is an optimization problem

$$\begin{aligned} \inf_{X \in \mathbb{S}^{n \times n}} \quad & \operatorname{tr}(CX) \\ \text{s.t.} \quad & \operatorname{tr}(A_i X) \le b_i, \quad i \in \{1, \ldots, m\} \\ & X \succeq 0, \end{aligned} \tag{SDP}$$

where $A_1, \ldots, A_m \in \mathbb{S}^{n \times n}$, $b_1, \ldots, b_m \in \mathbb{R}$ and $C \in \mathbb{S}^{n \times n}$.

Linear and semidefinite programs are well-studied classes of optimization problems, and there exist algorithms that can solve them efficiently to arbitrary accuracy. Notable examples include the *simplex method*, the *ellipsoid method* and various types of *interior-point methods*. We do not study these algorithms in this thesis.

For every conic program, we can define a dual problem.

Definition A.15 (Dual problem)**.** The *dual problem* to problem (P) is defined as

$$\begin{aligned} \sup_{y_1, \ldots, y_m \in \mathbb{R}} \quad & \sum_{i=1}^{n} y_i b_i \\ \text{s.t.} \quad & C - \sum_{i=1}^{m} y_i A_i \in \mathcal{K}^* \\ & y_i \le 0, \quad i \in \{1, \ldots, m\}. \end{aligned} \tag{D}$$

The dual problem of an LP or an SDP follows directly from this definition.

Theorem A.13. The dual problem to problem (LP) is given by

$$\begin{aligned} \sup_{y \in \mathbb{R}^m} \quad & y^\top b \\ \text{s.t.} \quad & y^\top A \le c \\ & y \le 0. \end{aligned} \tag{LP-D}$$

Theorem A.14. The dual problem to problem (SDP) is given by

$$\begin{aligned} \sup_{y_1,\ldots,y_m\in\mathbb{R}} \quad & \sum_{i=1}^{n} y_i b_i \\ \text{s. t.} \quad & \sum_{i=1}^{m} y_i A_i \preceq C \\ & y_i \leq 0, \quad i \in \{1,\ldots,m\}\,. \end{aligned} \tag{SDP-D}$$

It now follows that every feasible solution value of (D) is a lower bound on the feasible solution values of (P).

Theorem A.15. Let X be a feasible solution to problem (P), and $(y_1,\ldots,y_m)$ a feasible solution to problem (D). Then,

$$\langle C, X\rangle \geq \sum_{i=1}^{m} y_i b_i.$$

Proof. Since $X \in \mathcal{K}$ and $C - \sum_{i=1}^{m} y_i A_i \in \mathcal{K}^*$, we have

$$0 \leq \left\langle C - \sum_{i=1}^{m} y_i A_i, X \right\rangle = \langle C, X\rangle - \sum_{i=1}^{m} y_i \langle A_i, X\rangle \leq \langle C, X\rangle - \sum_{i=1}^{m} y_i b_i.$$

Hence, $\langle C, X\rangle \geq \sum_{i=1}^{m} y_i b_i$. $\square$

Corollary A.15.1 (Weak duality)**.** Let p^* be the optimal solution value of problem (P), and d^* the optimal solution value of problem (D) Then,

$$p^* \geq d^*.$$

A.5 Kronecker products

The matrix consisting of the elements of a matrix $A \in \mathbb{R}^{m \times n}$ multiplied by the elements of $B \in \mathbb{R}^{p \times q}$ is called the *Kronecker product* of A and B, and is denoted by $A \otimes B$.

Definition A.16 (Kronecker product)**.** Let

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n} \text{ and } B = \begin{bmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & & \vdots \\ b_{p1} & \cdots & b_{pq} \end{bmatrix} \in \mathbb{R}^{p \times q}.$$

Then, the *Kronecker product* $A \otimes B$ of A and B is defined as

$$\begin{aligned} A \otimes B &= \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} & \cdots & a_{11}b_{1q} & \cdots & a_{1n}b_{11} & \cdots & a_{1n}b_{1q} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{11}b_{p1} & \cdots & a_{11}b_{pq} & \cdots & a_{1n}b_{p1} & \cdots & a_{1n}b_{pq} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}b_{11} & \cdots & a_{m1}b_{1q} & \cdots & a_{mn}b_{11} & \cdots & a_{mn}b_{1q} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}b_{p1} & \cdots & a_{m1}b_{pq} & \cdots & a_{mn}b_{p1} & \cdots & a_{mn}b_{pq} \end{bmatrix}. \end{aligned}$$

The the matrix product and the Kronecker product satisfy a mixed-product property.

Theorem A.16 (Mixed-product property)**.** For all $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$, $C \in \mathbb{R}^{n \times r}$ and $D \in \mathbb{R}^{q \times s}$, we have

$$(A \otimes B)\,(C \otimes D) = AC \otimes BD.$$

Proof. Let

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n},\ B \in \mathbb{R}^{p \times q},\ C = \begin{bmatrix} c_{11} & \cdots & c_{1r} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nr} \end{bmatrix} \in \mathbb{R}^{n \times r}, \text{ and } D \in \mathbb{R}^{q \times s}.$$

Then,

$$
\begin{aligned}
(A \otimes B)(C \otimes D) &= \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \begin{bmatrix} c_{11}D & \cdots & c_{1r}D \\ \vdots & & \vdots \\ c_{n1}D & \cdots & c_{nr}D \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^{n} a_{1i}c_{i1}BD & \cdots & \sum_{i=1}^{n} a_{1i}c_{ir}BD \\ \vdots & & \vdots \\ \sum_{i=1}^{n} a_{mi}c_{i1}BD & \cdots & \sum_{i=1}^{n} a_{mi}c_{ir}BD \end{bmatrix} = AC \otimes BD
\end{aligned}
$$

□

The following theorem now states that the eigenvalues of $A \otimes B$ are the eigenvalues of A multiplied by the eigenvalues of B.

Theorem A.17. Let $A \in \mathbb{R}^{m \times m}$ with eigenvalues $\lambda_1, \ldots, \lambda_m$ and $B \in \mathbb{R}^{n \times n}$ with eigenvalues $\mu_1, \ldots, \mu_n$. Then, the eigenvalues of $A \otimes B$ are $\lambda_1\mu_1, \ldots, \lambda_1\mu_n, \ldots, \lambda_m\mu_1, \ldots, \lambda_m\mu_n$.

Proof. Let λ be an eigenvalue of A with an eigenvector x, and μ an eigenvalue of B with an eigenvector y. From the mixed-product property, it then follows that

$$
(A \otimes B)(x \otimes y) = Ax \otimes Ay = \lambda x \otimes \mu y = \lambda\mu (x \otimes y).
$$

Hence, $\lambda\mu$ is an eigenvalue of $A \otimes B$. □

Corollary A.17.1. Let $A \succ 0$ and $B \succ 0$. Then, $A \otimes B \succ 0$.

A.6 Schur complements

When row-reducing a block matrix $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}$, where A_{11} is invertible, we obtain

$$
\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \sim \begin{bmatrix} I_m & A_{11}^{-1}A_{12} \\ A_{21} & A_{22} \end{bmatrix} \sim \begin{bmatrix} I_m & A_{11}^{-1}A_{12} \\ 0 & A_{22} - A_{22}A_{11}^{-1}A_{12} \end{bmatrix}.
$$

Here, the matrix $A_{22} - A_{22}A_{11}^{-1}A_{12}$ is called the *Schur complement* of the matrix A_{11}.

Definition A.17 (Schur complement)**.** Let

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}.$$

Assume that A_{11} is invertible. Then, $A_{22} - A_{21}A_{11}^{-1}A_{12}$ is called the *Schur complement* of A_{11}.

The following theorem shows that the block matrix $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is similar to the block diagonal matrix $\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$, and can be directly verified.

Theorem A.18. Let

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}.$$

Assume that A_{11} is invertible. Then,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ A_{21}A_{11}^{-1} & I_n \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I_m & A_{11}^{-1}A_{12} \\ 0 & I_n \end{bmatrix}.$$

It now follows directly that a block matrix

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{S}^{(m+n)\times(m+n)}$$

is positive-definite if and only if the matrix A_{11} and its Schur complement $A_{22} - A_{21}A_{11}^{-1}A_{12}$ are positive-definite.

Theorem A.19. Let

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{S}^{(m+n)\times(m+n)}.$$

Then, $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \succ 0$ if and only if $A_{11} \succ 0$ and $A_{22} \succ A_{21}A_{11}^{-1}A_{12}$.

Proof. $\Longleftarrow$) Assume that $A_{11} \succ 0$ and $A_{22} \succ A_{21}A_{11}^{-1}A_{12}$. Then,

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ A_{21}A_{11}^{-1} & I_n \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I_m & A_{11}^{-1}A_{12} \\ 0 & I_n \end{bmatrix} \succ 0.$$

$\Longrightarrow$) Assume that $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \succ 0$. Then, $A_{11} \succ 0$, and

$$\begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ -A_{21}A_{11}^{-1} & I_n \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_m & -A_{11}^{-1}A_{12} \\ 0 & I_n \end{bmatrix} \succ 0,$$

so also $A_{22} \succ A_{21}A_{11}^{-1}A_{12}$.

$\square$

Appendix B

Convergence of gradient descent and Newton's method

B.1 Gradient descent

Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}^n$ is a continuously differentiable function on an open domain $D \subseteq \mathbb{R}^n$.

We assume that the gradient ∇f is *Lipschitz continuous*.

Definition B.1 (Lipschitz continuous function)**.** Let V and W be normed vector spaces. Then, a function $f : D \to W$ on an open domain $D \subseteq V$ is called *Lipschitz continuous* if there exists an $L \geq 0$ such that, for all $x, y \in D$, we have

$$\| f(y) - f(x) \| \leq L \| y - x \| .$$

It is easy to show that every Lipschitz continuous function is continuous, hence the name. However, not every continuous function is also Lipschitz continuous.

To determine when gradient descent converges, we need the following lemma.

Lemma B.1 (Descent lemma; e.g., Beck, 2014, Lemma 4.22)**.** Let $f : D \to \mathbb{R}$ be a differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that ∇f is Lipschitz continuous with Lipschitz constant L. For all $x, y \in D$, we then have

$$f(y) \leq f(x) + (\nabla f(x))^\top (y - x) + \frac{L}{2} \| y - x \|^2 .$$

Proof. Let $x, y \in D$. From the chain rule and the fundamental theorem of calculus, it then follows that

$$f(y) - f(x) = \int_0^1 \left(\nabla f(x + t(y - x))\right)^\top (y - x)\, dt.$$

Therefore,

$$\begin{aligned}
\left|f(y) - f(x) - (\nabla f(x))^\top (y - x)\right| &= \left|\int_0^1 \left(\nabla f(x + t(y - x))\right)^\top (y - x)\, dt\right| \\
&\leq \int_0^1 \left|\left(\nabla f(x + t(y - x))\right)^\top (y - x)\right| dt \\
&\leq \int_0^1 \|\nabla f(x + t(y - x)) - \nabla f(x)\| \, \|y - x\|\, dt \\
&\leq \int_0^1 tL \|y - x\|^2\, dt = \frac{L}{2} \|y - x\|^2,
\end{aligned}$$

where the first inequality follows from the triangle inequality, the second from the Cauchy-Schwarz inequality, and the third from the Lipschitz property. Hence,

$$f(y) \leq f(x) + (\nabla f(x))^\top (y - x) + \frac{L}{2} \|y - x\|^2.$$

□

The following lemma now shows how much the value of the f after a single iteration.

Lemma B.2 (Sufficient-decrease lemma; e.g., Beck, 2014, Lemma 4.23)**.** Let $f : D \to \mathbb{R}$ be a differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that ∇f is Lipschitz continuous with Lipschitz constant L. For all $x \in D$, and $t > 0$ with $x - t\nabla f(x) \in D$, we then have

$$f(x) - f(x - t\nabla f(x)) \geq t\left(1 - \frac{tL}{2}\right) \|\nabla f(x)\|^2.$$

Proof. Let $x \in D$, and $t > 0$ with $x - t\nabla f(x) \in D$. From the descent lemma, it then follows that

$$\begin{aligned}
f(x - t\nabla f(x)) &\leq f(x) + (\nabla f(x))^\top (-t\nabla f(x)) + \frac{L}{2} \|-t\nabla f(x)\|^2 \\
&= f(x) - t\|\nabla f(x)\|^2 + \frac{t^2 L}{2} \|\nabla f(x)\|^2 \\
&= f(x) - t\left(1 - \frac{tL}{2}\right) \|\nabla f(x)\|^2.
\end{aligned}$$

Hence,

$$f(x) - f(x - t\nabla f(x)) \geq t\left(1 - \frac{tL}{2}\right) \|\nabla f(x)\|^2.$$

□

It now follows that gradient descent converges if the step size is sufficiently small.

Theorem B.3 (Convergence of gradient descent; e.g., Beck, 2014, Theorem 4.24 and Theorem 4.25)**.** Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}$ is a continuously differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that f is bounded from below, and that ∇f is Lipschitz continuous with Lipschitz constant $L > 0$. For every sequence $(x_k)_{k=0}^{\infty}$ obtained by gradient descent with step size $t < \frac{2}{L}$, we then have

$$\nabla f(x_k) \to 0 \text{ as } k \to \infty.$$

Proof. Let $(x_k)_{k=0}^{\infty}$ be a sequence obtained by gradient descent with step size $t < \frac{2}{L}$. From the sufficient-decrease lemma, it then follows that, for all $k \in \mathbb{N}_0$, we have

$$f(x_k) - f(x_{k+1}) \geq t\left(1 - \frac{tL}{2}\right) \|\nabla f(x_k)\|^2 .$$

Since $0 < t < \frac{2}{L}$, we then have $t\left(1 - \frac{tL}{2}\right) > 0$. Therefore, for all $k \in \mathbb{N}_0$, we have

$$f(x_k) - f(x_{k+1}) \geq t\left(1 - \frac{tL}{2}\right) \|\nabla f(x_k)\|^2 \geq 0,$$

so $f(x_{k+1}) \leq f(x_k)$, and $f(x_{k+1}) = f(x_k)$ if and only if $\nabla f(x_k) = 0$. Hence, the sequence $(f(x_k))_{k=0}^{\infty}$ is nonincreasing. Since it is bounded from below, it is convergent. Therefore,

$$0 \leq t\left(1 - \frac{tL}{2}\right) \|\nabla f(x_k)\|^2 \leq f(x_k) - f(x_{k+1}) \to 0 \text{ as } k \to \infty.$$

Hence, $\nabla f(x_k) \to 0$ as $k \to \infty$. □

B.2 Newton's method

Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}^n$ is a twice continuously differentiable function on an open domain $D \subseteq \mathbb{R}^n$.

We assume that the Hessian $\nabla^2 f$ is Lipschitz continuous. Then, Newton's method converges to a local minimizer x^* of f if the initial point is sufficiently close to x^*.

Theorem B.4 (Local quadratic convergence of Newton's method; e.g., Beck, 2014, Theorem 5.2)**.** Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}^n$ is a twice continuously differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that f has a local minimizer x^* with $\nabla^2 f(x^*) \succ 0$, and that $\nabla^2 f$ is Lipschitz continuous with Lipschitz constant $L > 0$. Then, there exist an $r > 0$ and a $c > 0$ such that, for every sequence $(x_k)_{k=0}^{\infty}$ obtained by Newton's method with $\|x_0 - x^*\| \leq r$, we have

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^2$$

for all $k \in \mathbb{N}_0$.

Proof. Since D is an open set, and $\nabla^2 f(x^*) \succ 0$, there exist an $m > 0$ and an $r' > 0$ such that, for all $x \in \mathbb{R}^n$, we have $x \in D$ with $\|x - x^*\| \leq r'$, and $\nabla^2 f(x) \succeq m I_n$.

Let $(x_k)_{k=0}^{\infty}$ be a sequence obtained by the d^{th}-order Newton method, and let $k \in \mathbb{N}_0$. Assume that $\|x_k - x^*\| \leq r'$. Since x^* is a local minimizer of f, we have $\nabla f(x^*) = 0$. From the chain rule and the fundamental theorem, it now follows that

$$\begin{aligned}
x_{k+1} - x^* &= x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k) - x^* \\
&= x_k - x^* - \left(\nabla^2 f(x_k)\right)^{-1} \left(\nabla f(x_k) - \nabla f(x^*)\right) \\
&= x_k - x^* - \left(\nabla^2 f(x_k)\right)^{-1} \int_0^1 \nabla^2 f(x^* + t(x_k - x^*))(x_k - x^*)\, dt \\
&= \left(\nabla^2 f(x_k)\right)^{-1} \int_0^1 \left(\nabla^2 f(x_k) - \nabla^2 f(x^* + t(x_k - x^*))\right)(x_k - x^*)\, dt.
\end{aligned}$$

Since $\nabla^2 f(x_k) \succeq m I_n$, we have $\left(\nabla^2 f(x_k)\right)^{-1} \preceq \frac{1}{m} I_n$. Hence,

$$\begin{aligned}
\|x_{k+1} - x^*\| &= \left\| \left(\nabla^2 f(x_k)\right)^{-1} \int_0^1 \left(\nabla^2 f(x_k) - \nabla^2 f(x^* + t(x_k - x^*))\right)(x_k - x^*)\, dt \right\| \\
&\leq \left\| \left(\nabla^2 f(x_k)\right)^{-1} \right\| \left\| \int_0^1 \left(\nabla^2 f(x_k) - \nabla^2 f(x^* + t(x_k - x^*))\right)(x_k - x^*)\, dt \right\| \\
&\leq \left\| \left(\nabla^2 f(x_k)\right)^{-1} \right\| \int_0^1 \left\| \left(\nabla^2 f(x_k) - \nabla^2 f(x^* + t(x_k - x^*))\right)(x_k - x^*) \right\| dt \\
&\leq \left\| \left(\nabla^2 f(x_k)\right)^{-1} \right\| \int_0^1 \left\| \nabla^2 f(x_k) - \nabla^2 f(x^* + t(x_k - x^*)) \right\| \|x_k - x^*\|\, dt \\
&\leq \frac{1}{m} \int_0^1 (1 - t) L \|x_k - x^*\|^2\, dt = \frac{L}{2m} \|x_k - x^*\|^2,
\end{aligned}$$

where first and third inequalities follow from Theorem A.4, the second from the triangle inequality, and the fourth from the Lipschitz property.

Let $r = \min\left\{r', \frac{2m}{L}\right\}$ and $c = \frac{L}{2m}$. Then, $r \leq \frac{1}{c}$, so $cr^2 \leq r$.

Let $(x_k)_{k=0}^{\infty}$ be a sequence obtained by Newton's method with $\|x_0 - x^*\| \leq r$. If $\|x_k - x^*\| \leq r$ for some $k \in \mathbb{N}_0$, then

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^2 \leq cr^2 \leq r.$$

Hence, $\|x_k - x^*\| \leq r$, so $\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^d$, for all $k \in \mathbb{N}_0$. □

Corollary B.4.1 (e.g., Beck, 2014, Theorem 5.2)**.** Consider an unconstrained optimization problem

$$\inf_{x \in D} \quad f(x),$$

where $f : D \to \mathbb{R}^n$ is a twice continuously differentiable function on an open convex domain $D \subseteq \mathbb{R}^n$. Assume that f has a local minimizer x^* with $\nabla^2 f(x^*) \succ 0$, and that $\nabla^2 f$ is Lipschitz continuous with Lipschitz constant $L > 0$. Then, there exists an $R > 0$ such that, for every sequence $(x_k)_{k=0}^{\infty}$ obtained by Newton's method with $\|x_0 - x^*\| \leq R$, we have

$$\|x_k - x^*\| \leq 2R \left(\frac{1}{2}\right)^{2^k}$$

for all $k \in \mathbb{N}_0$.

Proof. From local quadratic convergence of Newton's method, it follows that there exist an $r > 0$ and a $c > 0$ such that, for every sequence $(x_k)_{k=0}^{\infty}$ obtained by Newton's method with $\|x_0 - x^*\| \leq r$, we have $\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^2$ for all $k \in \mathbb{N}_0$.

Let $R = \min\left\{r, \frac{1}{2c}\right\}$. Then, $2R \leq \frac{1}{c}$, so $c(2R)^2 \leq 2R$.

Let $(x_k)_{k=0}^{\infty}$ be a sequence obtained by Newton's method with $\|x_0 - x^*\| \leq R$. If $\|x_k - x^*\| \leq 2R \left(\frac{1}{2}\right)^{2^k}$ for some $k \in \mathbb{N}_0$, then

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^2 \leq c \left(2R \left(\frac{1}{2}\right)^{2^k}\right)^2 = c(2R)^2 \left(\frac{1}{2}\right)^{2^{k+1}} = 2R \left(\frac{1}{2}\right)^{2^{k+1}}.$$

Hence, $\|x_k - x^*\| \leq 2R \left(\frac{1}{2}\right)^{2^k}$ for all $k \in \mathbb{N}_0$. □

Bibliography

Ahmadi, A. A., Chaudhry, A., and Zhang, J. (2024). Higher-Order Newton Methods with Polynomial Work per Iteration. *Advances in Mathematics*, 452:109808.

Ahmadi, A. A. and Hall, G. (2018). DC Decomposition of Nonconvex Polynomials with Algebraic Techniques. *Mathematical Programming*, 169(1):69–94.

Barvinok, A. (2002). *A Course in Convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.

Beck, A. (2014). *Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB*. SIAM, Philadelphia, PA, 2 edition.

Bellare, M. and Rogaway, P. (1995). The Complexity of Approximating a Nonlinear Program. *Mathematical Programming*, 69(3):429–441.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, Cambridge, UK.

Clenshaw, C. and Curtis, A. (1960). A Method for Numerical Integration on an Automatic Computer. *Numerische Mathematik*, 2(1):197–205.

De Klerk, E. and Laurent, M. (2011). On the Lasserre Hierarchy of Semidefinite Programming Relaxations of Convex Polynomial Optimization Problems. *SIAM Journal on Optimization*, 21(3):824–832.

Doikov, N. and Nesterov, Y. (2021). Local Convergence of Tensor Methods. In *Mathematical Optimization Theory and Operations Research (MOTOR 2021)*, volume 12755 of *Lecture Notes in Computer Science*, pages 134–150. Springer.

Doikov, N. and Nesterov, Y. (2024). High-Order Optimization Methods for Fully Composite Problems. *Mathematical Programming*.

Folland, G. B. (2002). *Advanced Calculus*. Prentice Hall, Upper Saddle River, NJ, 1 edition.

Helton, J. W. and Nie, J. (2010). Semidefinite Representation of Convex Sets. *Mathematical Programming*, 122(1):21–64.

Lasserre, J. B. (2001). Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, 11(3):796–817.

Lasserre, J. B. (2009). Convexity in Semialgebraic Geometry and Polynomial Optimization. *SIAM Journal on Optimization*, 19(4):1995–2021.

Lay, D. C., Lay, S. R., and McDonald, J. J. (2021). *Linear Algebra and Its Applications*. Pearson, 6th edition.

Nesterov, Y. and Polyak, B. T. (2006). Cubic regularization of newton method and its global performance. *Mathematical Programming*, 108(1):177–205.

Saint Raymond, X. (1991). *Elementary Introduction to the Theory of Pseudodifferential Operators*. Studies in Advanced Mathematics, CRC Press, Boca Raton.

Zhu, W. and Cartis, C. (2024). Global Convergence of High-Order Regularization Methods with Sums-of-Squares Taylor Models. `https://arxiv.org/abs/2406.03863`. arXiv preprint arXiv:2406.03863.