Notes on Constrained Optimization

Wes Cowan
Department of Mathematics, Rutgers University
110 Frelinghuysen Rd., Piscataway, NJ 08854

December 16, 2016

1 Introduction

In the previous set of notes, we considered the problem of unconstrained optimization, minimization of a scalar function $f(\underline{x})$ over all $\underline{x} \in \mathbb{R}^n$. In these notes, we consider the problem of *constrained* optimization, in which the set of feasible $\underline{x}$ is restricted. That is, given a function $f : \mathbb{R}^n \mapsto \mathbb{R}$, solve the following problem:

$$\begin{aligned} &\text{minimize } f(\underline{x}) \\ &\text{such that (s.t.) } \underline{x} \in X \\ &X \subset \mathbb{R}^n, \end{aligned} \tag{1}$$

where X is taken to be a known or given feasible set. In particular, we are interested in identifying $\underline{x}^* \in X$ such that $f(\underline{x}^*)$ realizes the above minimum, if such an $\underline{x}^*$ exists. Typical constraints include bounds on $\underline{x}$, such as $X = \{\underline{x} \in \mathbb{R}^n : ||\underline{x}|| \leqslant R\}$ for some $R > 0$, or linear inequality constraints such as $X = \{\underline{x} \in \mathbb{R}^n : A\underline{x} \geqslant \underline{b}\}$ for a matrix A and a column vector $\underline{b}$. More examples will be discussed.

In the usual way, we subsume the problem of maximization of f by instead minimizing $-f$. As in the previous notes, unless otherwise indicated we take f to be continuous, and at least twice differentiable over the domain. We additionally will largely restrict ourselves to constraint sets X that are *closed*; this will ensure that if f has a minimum value over X, it is realized by some point in the constraint set X. (Consider, for instance, trying to maximize the function x over the open interval $X = (0,1)$.) This will not be a serious limitation in practice.

We may draw a distinction between *constrained local minima* and *constrained global minima*:

Definition 1 *A **constrained local minimum** occurs at $\underline{x}^* \in X$ if there exists an $\varepsilon > 0$ such that*

$$f(\underline{x}) \geqslant f(\underline{x}^*) \text{ for all } \underline{x} \in X \text{ with } ||\underline{x} - \underline{x}^*|| < \varepsilon. \tag{2}$$

*A **constrained global minimum** occurs at $\underline{x}^* \in X$ if*

$$f(\underline{x}) \geqslant f(\underline{x}^*) \text{ for all } \underline{x} \in X. \tag{3}$$

We say that a minimum is strict if in the above inequalities we have $f(\underline{x}) > f(\underline{x}^*)$ for $\underline{x} \neq \underline{x}^*$.

As in the case of unconstrained optimization, it may be difficult to determine whether a given local minimum over X is in fact a global minimum over X. Much will depend, for instance, on the structure or geometry of X as well as the properties of f (such as convexity). As in the case of unconstrained optimization, we will largely focus on the following questions:

- When is a minimum of f over X, and corresponding minimizer $\underline{x}^* \in X$, guaranteed to exist?
- How can minimizers be accurately and efficiently computed?
- What are the benefits and costs of various algorithmic solutions?
- What kind of theoretical guarantees do the various algorithms offer?

It is instructive to review the established framework and methodology for unconstrained optimization, to see how it must be modified in the case of constrained optimization. In the unconstrained case, minima were identified by considering stationary points where $\nabla f(\underline{x}^*) = 0$, which provided a system of (potentially non-linear) equations that could be solved for $\underline{x}^*$. In the constrained case, however, it can occur that the minimum of f over a restricted set X is not a minimum of f over $\mathbb{R}^n$, and hence is not a stationary point of f. This requires the derivation of different, though related, characterizations of minima in the restricted case. Similarly, in the unconstrained case, algorithms for computing minimizers proceeded by generating a series of iterates $\{\underline{x}_k\}$ starting from an initial guess and successively moving in directions of descent of f. In the constrained case, we must be careful that motion in a given direction does not carry us out of the constraint set X - these directions are said to be *unfeasible*. The case of constrained optimization will require new analytic and algorithmic tools to extend our previous results from the unconstrained case.

Speaking broadly, there are two main ways to approach these problems: the geometric perspective which views X as a geometric object existing in $\mathbb{R}^n$ (for instance, X as a sphere, box, or convex set); and the analytic perspective which defines X in terms of a set of equalities and inequalities that any $\underline{x} \in X$ must satisfy. We will examine the problem from both these angles, and derive corresponding minimization algorithms.

This set of notes may at times reference the previous set of notes on Unconstrained Optimization for necessary background material - revisit those notes as necessary. In this set of notes, Section 2 gives a number of motivating examples of constrained optimization problems, and section 3 a number of examples of possible constraint sets of interest, including a brief discussion of the important case of linear inequality constraints or X as convex polytopes (a generalization of polyhedra). Section 4 analyzes the problem of existence and characterization of minimizers from the geometric perspective on X, heavily utilizing convex analysis. This line of thought is continued in section 5, which utilizes the geometric perspective to derive descent-type algorithms analogous to the descent algorithms for unconstrained optimization. Section 6 introduces the analytic perspective, defining X in terms of a system of equalities or inequalities, and presents results on the existence and characterization of minima in this case, namely the central result of the existence of Lagrange multipliers. This analytic perspective is continued in section 7, which derives new minimization algorithms based on the analytic perspective. Lagrange multipliers and the Lagrangian lead to the important concept of duality, which provides an important theoretical understanding of the structure of optimization problems. Duality is explored in section 8. This leads to a number of new minimization algorithms, discussed in section **??**.

2 Some Motivating Discussion

Constrained optimization arises in a variety of contexts. Two frequent examples in practice can be thought of in the following way: let $\underline{x}$ be a vector of parameters governing some process, decision, or business (for instance, inventory to buy or quantities to produce). A classical problem then is to find $\underline{x}^* \in \mathbb{R}^n$ to realize

$$\begin{aligned} \text{maximize } & \text{PerformanceUnder}(\underline{x}) \\ \text{s.t. } & \text{CostUnder}(\underline{x}) \leqslant \text{budget}, \end{aligned} \tag{4}$$

where PerformanceUnder($\underline{x}$) represents the response, potentially profits or rewards, for a given set of parameters, and CostUnder($\underline{x}$) represents costs (for instance, energy usage or costs of materials) associated with acting under a given set of parameters, which is limited by some specified budget. In short, maximize performance subject to cost constraints.

A frequent dual problem to the above is to minimize cost while guaranteeing a sufficient level of performance, i.e.,

$$\begin{aligned} \text{minimize } & \text{CostUnder}(\underline{x}) \\ \text{s.t. } & \text{PerformanceUnder}(\underline{x}) \geqslant \text{minPerformance}. \end{aligned} \tag{5}$$

The specific forms of the Performance or Cost functions will depend on the specific problem being addressed.

An Illustration via Linear Programming: A classic example of this type of problem is linear programming, in which the objective function f is a linear function of the variables, and variables are constrained by a set of linear inequalities. A traditional application is as follows: for a given $\underline{x} \in \mathbb{R}^n$, let x_i represent a quantity of product i to be produced, to be sold at (known) cost c_i, for $i = 1, \ldots, n$. In this case, the return on any production schedule $\underline{x}$ is given by

$$c_1 x_1 + c_2 x_2 + \ldots + c_n x_n = \underline{c}^{\mathrm{T}} \underline{x}. \tag{6}$$

However there may be constraints on how much of product i can be produced. Suppose that each product i is composed of various materials $j = 1, \ldots, m$. Let $a_{j,i}$ be the proportion of material j needed to make product i, and let b_j be the total amount of material j available. In this case, a given production schedule $\underline{x}$ is constrained by

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n &\geqslant b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,n}x_n &\geqslant b_2 \\ &\ldots \\ a_{m,1}x_1 + a_{m,2}x_2 + \ldots + a_{m,n}x_n &\geqslant b_m, \end{aligned} \tag{7}$$

or equivalently, defing the matrix $A = [a_{j,i}]_{m \times n}$ and the budget vector $\underline{b}$,

$$A\underline{x} \geqslant \underline{b}. \tag{8}$$

Taking the constraint as well that under any production schedule $\underline{x}$, the amount of product i must be non-negative, the problem of optimizing the production schedule may be formulated as

$$\begin{aligned} \text{maximize } & \underline{c}^{\mathrm{T}} \underline{x} \\ \text{s.t. } & A\underline{x} \geqslant \underline{b} \\ & x_i \geqslant 0 \text{ for } i = 1, 2, \ldots, n. \end{aligned} \tag{9}$$

Note, the positivity constraints on the x_i may be absorbed into the other linear constraints by appropriately expanding the matrix A and budget vector $\underline{b}$.

The field of linear programming is broad and well established, with many results and algorithms on the existence and determination of optima. In particular, it can be shown feasible sets defined by linear inequalities of the variables define (potentially unbounded) polytopes in the space of $\mathbb{R}^n$, and that the optima of the objective function (if they exist) must lie on the vertices of this polytope. This can, to some measure, reduce the problem to systematically checking the value of the objective function at each of the (finitely many) vertices of the feasible polytope, i.e., the simplex algorithm.

In these notes and in this class, we are interested in the more general problem of *non-linear programming*, when the objective function and the constraint set are either or both defined in terms of non-linear functions. These problems have a much richer structure than linear programs. However, it is worth observing that in the world of model-building and optimization, linear models are frequently a good and useful tool, especially as a starting place for building more complex models.

2.1 Examples of Constrained Problems

In this section, we consider a few interesting examples of non-linear programming problems.

Maximum Likelihood Estimation:
In many contexts, collected data is used to estimate the parameters of some underlying model. This frequently takes the form of *maximum likelihood estimation* when the underlying model is probabilistic or random. As an example of this, consider taking a sequence of random, independent observations, where outcome i is observed with some unknown probability p_i, for $i = 1, \ldots, n$ (the classic example here is a sequence of coin flips). Suppose that N observations are taken, and outcome i is observed N_i times (so $N_1 + \ldots + N_n = N$). In that case, if the true underlying probabilities are given by the vector $\underline{p}$, then the probability of observing the data collected will be

$$\prod_{i=1}^{n} p_i^{N_i}. \tag{10}$$

If the true probabilities are unknown, we may wish to estimate them based on the collected data. The technique of maximum likelihood estimation is to find the set of parameters $\underline{p}^*$ that maximizes the likelihood of the observed data. Observing the constraints on feasible probabilities - they must sum to 1 and be non-negative - we have the following non-linear program:

$$\begin{aligned} \text{maximize } & \prod_{i=1}^{n} p_i^{N_i} \\ \text{s.t. } & p_1 + \ldots + p_n = 1 \\ & p_i \geqslant 0 \text{ for } i = 1, 2, \ldots, n. \end{aligned} \tag{11}$$

An equivalent and potentially simpler formulation is to minimize the $-\ln$ of the above objective function, reducing the product to a sum, albeit one still non-linear as a function of the relevant variables (though with the caveat that we are taking $0 * \ln 0$ to be 0):

$$\begin{aligned} \text{minimize } & -\sum_{i=1}^{n} N_i \ln(p_i) \\ \text{s.t. } & p_1 + \ldots + p_n = 1 \\ & p_i \geqslant 0 \text{ for } i = 1, 2, \ldots, n. \end{aligned} \tag{12}$$

In either case, it can be shown that the optimal solution is to take $p_i^* = N_i/N$, i.e., the most likely probability is the observed frequency.

Defining a Representative:

Suppose a set of points have been specified, $\underline{y}_1, \ldots, \underline{y}_N$. These points may represent data, such as classified images or feature vectors. We wish to define a 'representative' point $\underline{x}^*$ that in some sense captures or summarizes this cluster of points, or characterizes some measure of similarity with all of them. There are a number of ways of defining such a point, for instance taking the average $(1/N)(\underline{y}_1 + \ldots + \underline{y}_N)$. In this case we wish to identify the $\underline{x}^*$ that minimizes the maximum distance to any of the $\underline{y}_i$. That is, $\underline{x}^*$ satisfies

$$\text{minimize}_{\underline{x}} \ \max\{||\underline{x} - \underline{y}_1||, ||\underline{x} - \underline{y}_2||, \ldots, ||\underline{x} - \underline{y}_N||\}. \tag{13}$$

As specified above, the problem is an unconstrained minimization problem. However, the objective function, as the maximum of N other functions, is potentially difficult to handle analytically, as the derivative may not be continuous. It can be massaged into the form of a constrained minimization problem though, in the following way, which will lend itself to the methods described in these notes:

$$\begin{aligned} \text{minimize}_{\underline{x},\delta} \ & \delta \\ \text{s.t. } & ||\underline{x} - \underline{y}_1|| \leqslant \delta \\ & ||\underline{x} - \underline{y}_2|| \leqslant \delta \\ & \ldots \\ & ||\underline{x} - \underline{y}_N|| \leqslant \delta. \end{aligned} \tag{14}$$

Underdetermined Linear Equations and Projection:

Let A be a matrix with m rows and n columns, and $\underline{b}$ be an m-dimensional vector. We take the rows of A to be linearly independent. If $n = m$, then A is invertible, and the system of equations

$$A\underline{x} = \underline{b} \tag{15}$$

has a unique solution given by $\underline{x}^* = A^{-1}\underline{b}$. If, however, $n > m$, then the system is underdetermined, and grants no unique solution. In such a case, a secondary criterion may be specified to determine the optimal solution. For instance, solutions $\underline{x}$ with smaller norms might be preferred. This leads to the constrained minimization problem which will in fact have a unique solution

$$\begin{aligned} \text{minimize } & ||\underline{x}||^2 \\ \text{s.t. } & A\underline{x} = \underline{b}. \end{aligned} \tag{16}$$

This is actually a specific instance of a more general problem, that of *projection*, which will be discussed in more detail. In particular, given a feasible set X (for instance, as specified by solutions to $A\underline{x} = \underline{b}$), we define the projection of a point $\underline{z}$ to be the point in X nearest to $\underline{z}$, i.e., solving

$$\begin{aligned} \text{minimize } & ||\underline{x} - \underline{z}||^2 \\ \text{s.t. } & \underline{x} \in X. \end{aligned} \tag{17}$$

The underdetermined linear equations example above can be seen as an example of this, taking $\underline{z} = 0$ and $X = \{\underline{x} : A\underline{x} = \underline{b}\}$.

2.1.1 An Example Problem

As an example, to motivate much of the analysis to follow, consider the following problem:

$$\begin{aligned} \text{maximize } & x_1 + x_2 + \ldots + x_n \\ \text{s.t. } & ||\underline{x}|| \leqslant R. \end{aligned} \tag{18}$$

Note, the constraint above effectively defines feasible $\underline{x}$ as restricted to a closed sphere of radius R, centered on the origin, i.e.,

$$x_1^2 + \ldots + x_n^2 \leqslant R^2. \tag{19}$$

We may consider the question of what an optimal solution $\underline{x}^*$ would 'look' like. For instance, for any optimal $\underline{x}^*$, all components must be non-negative. If $x_i^* < 0$ for some i, we could replace x_i^* with $-x_i^*$, which would leave the constraint satisfied, but increase the value of the objective function (by switching a negative term to a positive one).

Further, it must be that $||\underline{x}^*|| = R$, which is to say that the solution must lie on the boundary of the sphere. If not, if $||\underline{x}^*|| < R$, one of the components x_i may be increased without violating the constraint, which in turn increases the value of the objective function. This is an instance of a more general proposition to be discussed later: the maximum of a convex function over a closed convex set occurs on the boundary.

At this point, there are three ways analysis can proceed:

Reducing the Dimension of the Problem: Let $\underline{x}$ be some feasible point with $||\underline{x}|| = R$ and $x_i \geqslant 0$ for each i. Consider the following process for generating a new feasible $\underline{x}'$: let i and j be such that $x_i \neq x_j$; for $k \neq i$ and $k \neq j$, let $x'_k = x_k$, but let

$$x'_i = x'_j = \sqrt{\frac{x_i^2 + x_j^2}{2}}. \tag{20}$$

In short, construct $\underline{x}'$ by equalizing the i and j-th components, and then normalizing those components so that $||\underline{x}'|| = R$ again. It is straightforward algebra to show from this construction that

$$\sum_{k=1}^{n} x'_k > \sum_{k=1}^{n} x_k. \tag{21}$$

Hence, given any feasible $\underline{x}$ with two components that are not equal, we may construct a new feasible $\underline{x}'$ with a strictly larger value of the objective function. This is essentially a 1-dimensional ascent step over the feasible set. It follows then that any $\underline{x}$ may be improved upon (relative to the value of the objective function), *unless* all the components x_i are equal. Such a point would be stationary under this process, and indeed uniquely stationary. Hence we have a unique maximum, given by $\underline{x}^* = (R/\sqrt{n}, R/\sqrt{n}, \ldots, R/\sqrt{n})$.

Utilizing the Geometry of the Problem: Note that we may restate the original problem in the following form

$$\begin{aligned} \text{maximize } & \underline{u}^{\mathrm{T}}\underline{x} \\ \text{s.t. } & ||\underline{x}|| \leqslant R, \end{aligned} \tag{22}$$

where $\underline{u}$ is the vector of all 1's. Hence, we are simply trying to maximize the dot product with $\underline{u}$ over all vectors $\underline{x}$ with norm at most R. It is a classical result that a dot product is maximized by a vector in the same direction as the first, i.e., $\underline{x}^* = \alpha \underline{u}$ for some $\alpha > 0$. This reduces the problem to

$$\begin{aligned} \text{maximize}_{\alpha>0} \ & \alpha ||\underline{u}||^2 \\ \text{s.t. } & \alpha ||\underline{u}|| \leqslant R. \end{aligned} \tag{23}$$

This reduces the problem to a single dimension, and is maximized taking $\alpha = R/||\underline{u}|| = R/\sqrt{n}$, which yields a final solution of $\underline{x}^* = (R/\sqrt{n}, R/\sqrt{n}, \ldots, R/\sqrt{n})$.

The Problem of Feasible Directions: This is easily the most abstract of the three approaches, but it suggests the motivation for an extremely useful theoretical tool to be discussed later, Lagrange multipliers. It is convenient to denote the objective function $f(\underline{x}) = x_1 + \ldots + x_n$.

For a given $\underline{x}$ on the surface of the sphere of radius R centered at the origin, consider the plane tangent to the sphere at $\underline{x}$. This plane can be considered as the set of all vector $\underline{y}$ such that $\underline{y} - \underline{x}$ is orthogonal to $\underline{x}$, i.e., $\underline{x}^{\mathrm{T}}(\underline{y} - \underline{x}) = 0$.

Note that for any $\underline{x}'$ 'below' the plane, i.e., $\underline{x}^{\mathrm{T}}(\underline{x}' - \underline{x}) < 0$ or $\underline{x}'$ on the same side of the plane as the origin, $\underline{x}' - \underline{x}$ is a *feasible direction* in the sense that for sufficiently small $\alpha > 0$, $\underline{x} + \alpha(\underline{x}' - \underline{x})$ is a feasible point, i.e., $||\underline{x} + \alpha(\underline{x}' - \underline{x})|| \leqslant R$. If the directional derivative of f in that direction is positive, then the objective function can be increased by moving some distance in that direction. This implies that if $\nabla f(\underline{x})^{\mathrm{T}}(\underline{x}' - \underline{x}) > 0$ for any $\underline{x}'$ below the tangent plane at $\underline{x}$, then $\underline{x}$ cannot give the constrained maximum of f.

Hence, any potential maximizer $\underline{x}$ must satisfy $\nabla f(\underline{x})^{\mathrm{T}}(\underline{x}' - \underline{x}) \leqslant 0$ for any $\underline{x}'$ below the tangent plane, i.e, $\underline{x}^{\mathrm{T}}(\underline{x}' - \underline{x}) < 0$.

However, the set of all $\underline{x}'$ such that $\underline{x}^{\mathrm{T}}(\underline{x}' - \underline{x}) < 0$ represents half of $\mathbb{R}^n$, on one side of the plane passing through the point $\underline{x}$, normal to $\underline{x}$. The set of all $\underline{x}'$ such that $\nabla f(\underline{x})^{\mathrm{T}}(\underline{x}' - \underline{x}) \leqslant 0$ represents half of $\mathbb{R}^n$, on one side of the plane passing through the point $\underline{x}$, normal to $\nabla f(\underline{x})$. If any point $\underline{x}'$ in the half-space defined by the former plane must lie in the half-space defined by the latter, the two half-spaces must coincide, i.e., the dividing planes must be equivalent. This implies that the normal vectors that define the planes must be parallel.

Hence we see that for any maximizer $\underline{x}^*$, we must have that the vectors $\underline{x}^*$ and $\nabla f(\underline{x}^*)$ must be parallel, i.e.,

$$\nabla f(\underline{x}^*) = \lambda \underline{x}^*, \tag{24}$$

for some scalar value λ.

Notice, the above analysis holds for all functions f in this case - it can be used therefore to derive optimality conditions for any function defined over a sphere centered at the origin. In this specific case, we have that for all $\underline{x}$, $\nabla f(\underline{x}) = \underline{u}$ where $\underline{u}$ is the vector of all 1's. Hence we have that $\underline{x}^* = (1/\lambda, 1/\lambda, \ldots, 1/\lambda)$ for some scalar value $\lambda \neq 0$ (because $\nabla f(\underline{x})$ is a non-zero vector everywhere, it is impossible for $\lambda = 0$). Taking the additional constraint that $||\underline{x}^*|| = R$, we may solve for λ to yield a final solution of

$$\underline{x}^* = (R/\sqrt{n}, R/\sqrt{n}, \ldots, R/\sqrt{n}). \tag{25}$$

As will be discussed later, this is essentially the motivation behind Lagrange multipliers.

2.2 Example Problems

1 For the $\underline{x}'$ as constructed in Eq. (20), verify that $||\underline{x}'|| = R$ and verify Eq. (21) to show that $\underline{x}'$ is not only feasible, but increases the value of the objective function.

2 Viewing the construction in the 'Reducing the Dimension of the Problem' subsection above as an iterative algorithm for converging to the solution, what is the rate of convergence? How should each i, j for modification be chosen? *Hint: For a feasible $\underline{x}$ and corresponding constructed $\underline{x}'$, examine $f(\underline{x}') - f(\underline{x})$ and $||\underline{x}' - \underline{x}||$.*

3 Constraint Sets

There are a number of constraint sets $X \subset \mathbb{R}^n$ that frequently occur in practice, and it is worth trying to develop some intuition for them.

- **Polytopes via Linear Inequalities:** In many contexts, the feasible set X may be defined as the set of $\underline{x} \in \mathbb{R}^n$ that satisfy a finite number of linear inequalities. For $j = 1, \ldots, m$, we may be interested in $\underline{x}$ that satisfy the linear inequality

  $$a_{j,1}x_1 + a_{j,2}x_2 + \ldots + a_{j,n}x_n \geqslant b_j, \tag{26}$$

  for known $\{a_{j,i}\}, b_j$, or defining the matrix $A = [a_{j,i}]_{m \times n}$ and vector $\underline{b} \in \mathbb{R}^m$, we may write the full system of inequalities as

  $$A\underline{x} \geqslant \underline{b}. \tag{27}$$

  Observe that this is actually a fairly general model, subsuming other special cases: in the case that we want $a_{j,1}x_1 + \ldots + a_{j,n}x_n \leqslant b_j$, we may consider alternately $(-a_{j,1})x_1 + \ldots + (-a_{j,n})x_n \geqslant -b_j$. In the case that we want $a_{j,1}x_1 + \ldots + a_{j,n}x_n = b_j$, we may introduce two separate inequalities, $a_{j,1}x_1 + \ldots + a_{j,n}x_n \geqslant b_j$ and $(-a_{j,1})x_1 + \ldots + (-a_{j,n})x_n \geqslant -b_j$, and augment the matrix A accordingly.

  Taking the constraint set to be

  $$X = \{\underline{x} \in \mathbb{R}^n : A\underline{x} \geqslant \underline{b}\}, \tag{28}$$

if X is non-empty it actually has an important geometric interpretation, as a *convex polytope* in the space $\mathbb{R}^n$ (generalizing the concept of a polyhedron in $\mathbb{R}^3$). Any single linear inequality defines a 'half-space', half of $\mathbb{R}^n$. The planar faces of X correspond to surfaces where a single inequality constraint is satisfied with equality, $a_{j,1}x_1 + \ldots + a_{j,n}x_n = b_j$. The edges and vertices of X are the regions where multiple inequality constraints are satisfied with equality simultaneously.

The fundamental result of linear programming is that when the objective function f is linear, and the feasible set is defined by a set of linear inequalities as above, if there is an optimum (i.e., $\inf_{\underline{x} \in X} f(\underline{x}) > -\inf$), then the optimum is achieved at a vertex of the polytope. Because there are only finitely many vertices of such a polytope, the optimum can therefore be found by systematically checking the value of the function at the vertices - this is essentially the basis of the classical *simplex algorithm*.

- **Linear Spaces and Manifolds:** As a related constraint set, we frequently encounter sets of the form

$$X = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = \underline{b}\}, \tag{29}$$

for an $m \times n$ matrix A and m-dimensional vector $\underline{b}$. As noted above, this is technically subsumed in the inequality case (taking $A\underline{x} \geqslant \underline{b}$ and $-A\underline{x} \geqslant -\underline{b}$ simultaneously), but it is worth considering in its own right. We assume that A and $\underline{b}$ are such that X is not empty.

In particular, note that if $\underline{b} = 0$, X as above defines a *linear space*, for instance in $\mathbb{R}^3$ for various $A \neq 0$ the set X will give lines or planes passing through the origin.

Let $X_0 = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = 0\}$. For generate X, note that if $\underline{x}, \underline{y} \in X$, we have that $\underline{x} - \underline{y} \in X_0$. As such, if we specify a solution $\underline{x}'$ satisfying $A\underline{x}' = \underline{b}$, any point in X may be represented as $\underline{x}'$, plus some point in X_0. That is,

$$X = \{\underline{x}' + \underline{v} : \underline{v} \in X_0\}. \tag{30}$$

We see that X represents a *linear manifold*, a linear space translated to pass through the point $\underline{x}'$.

- **Spheres and Quadratic Surfaces:** A basic geometric constraint set of interest is the sphere, i.e., for some $\underline{x}_0$ and $R > 0$,

$$X = \{\underline{x} \in \mathbb{R}^n : ||\underline{x} - \underline{x}_0|| \leqslant R\}, \tag{31}$$

to define a sphere centered at $\underline{x}_0$ of radius R. We might also restrict to those vectors that are at distance *exactly* R from $\underline{x}_0$, i.e.,

$$X = \{\underline{x} \in \mathbb{R}^n : ||\underline{x} - \underline{x}_0|| = R\}. \tag{32}$$

This can be considered as a specific instance of a more general class of constraint sets, defined by quadratic surfaces.

Linear functions are frequently useful for objective functions or constraints if there is some notion of independence between the different components, e.g., one component can be modified without influencing or affecting the others. In terms of modeling, this is frequently a good starting assumption. However, when introducing elements such as covariance or dependence, the next model worth considering would be to include pairwise interactions, modeled by including terms such as $x_i x_j$. This leads generally to considering quadratic forms $\underline{x}^{\mathrm{T}} Q \underline{x} + \underline{c}^{\mathrm{T}} \underline{x}$ for a given matrix Q and vector $\underline{c}$, i.e.,

$$X_{\mathrm{P}} = \{\underline{x} \in \mathbb{R}^n : \underline{x}^{\mathrm{T}} Q \underline{x} + \underline{c}^{\mathrm{T}} \underline{x} \leqslant q\} \tag{33}$$

- **Simplices and Convex Hulls:** Consider defining the constraint set X_{P} to be the set of feasible probability vectors, i.e.,

  $$X_{\mathrm{P}} = \{\underline{p} \in \mathbb{R}^n : \sum_{i=1}^{n} p_i = 1 \text{ and } \forall i, p_i \geqslant 0\}. \tag{34}$$

  This is a special case of what is known as a *convex hull*. In the general case, consider a set of n points $\{\underline{v}_1, \ldots, \underline{v}_n\}$, and define the set X to be the set of all convex combinations of these points:

  $$X = \{\delta_1 \underline{v}_1 + \delta_2 \underline{v}_2 + \ldots + \delta_n \underline{v}_n : \sum_{i=1}^{n} \delta_i = 1 \text{ and } \forall i, \delta_i \geqslant 0\}. \tag{35}$$

  In this case, it can be shown that X is the smallest convex set that contains the points $\{\underline{v}_i\}$; this is the definition of a convex hull. Note, it can be shown that for an X defined by a system of linear inequalities as in the first case above, if X is bounded (i.e., can be contained in a sphere of finite radius), then X is the convex hull of its vertices.

  In the case of X_{P} as above, it can be shown that X_{P} is the convex hull of the n-many points $\{(1,0,\ldots,0),(0,1,\ldots,0),\ldots,(0,0,\ldots,1)\}$. The convex combinations of these points have natural interpretations as vectors of probabilities.

- **Orthants and Boxes:** The following is another special case of some of the examples considered above, but worth mentioning in its own right. In many cases, there are simple feasibility constraints on $\underline{x}$, for example taking $x_i \geqslant 0$ for all i. This defines the 'positive orthant' of $\mathbb{R}^n$ (generalizing the idea of a quadrant in $\mathbb{R}^2$),

  $$X = \{\underline{x} \in \mathbb{R}^n : \forall i, x_i \geqslant 0\}. \tag{36}$$

  Other orthants of $\mathbb{R}^n$ may be defined similarly. As a related concept, a useful constraint set is frequently to bound the individual components of $\underline{x}$, i.e., for a given set of constants $(a_1,b_1),\ldots,(a_n,b_n)$,

  $$X = \{\underline{x} \in \mathbb{R}^n : \forall i, a_i \leqslant x_i \leqslant b_i\}. \tag{37}$$

  This is, essentially, a high dimensional box.

- **General Equality and Inequality Constraints:** In many situations, including some of the above, the constraint set may be defined in terms of a system of equalities or inequalities that any feasible $\underline{x} \in \mathbb{R}^n$ must satisfy. In the most general case, we may take $h_1, \ldots, h_m$ as a set of functions $h_i : \mathbb{R}^n \mapsto \mathbb{R}$ and $g_1, \ldots, g_r$ as a set of functions $g_j : \mathbb{R}^n \mapsto \mathbb{R}$ and define

  $$X = \{\underline{x} \in \mathbb{R}^n : \forall i, h_i(\underline{x}) = 0 \text{ and } \forall j, g_j(\underline{x}) \leqslant 0\}. \tag{38}$$

It is important to observe that to some extent we may translate freely between equality constraints and inequality constraints. A given equality constraint $h_i(\underline{x}) = 0$ may be replaced by the two inequality constraints $h_i(\underline{x}) \leqslant 0$ and $-h_i(\underline{x}) \leqslant 0$. A given inequality constraints $g_j(\underline{x}) \leqslant 0$ may be replaced by an equality constraint by adding a variable x_{n+1} (increasing the space from $\mathbb{R}^n$ to $\mathbb{R}^{n+1}$) and taking the equality constraint $g_j(\underline{x}) + x_{n+1}^2 = 0$. The precise form of the constraints used may have some bearing on what computational methods will be effective; it is useful to be able to translate between them.

3.1 Example Problems

1 Prove that a polytope defined by a finite system of linear inequalities ($X = \{\underline{x} \in \mathbb{R}^n : A\underline{x} \geqslant \underline{b}\}$ as in the previous section) has finitely many vertices. *Hint: What do the vertices represent, relative to the system of inequalities?*

2 If $g : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex function, show that the set $X_g = \{\underline{x} \in \mathbb{R}^n : g(\underline{x}) \leqslant 0\}$ is a convex set.

3 Let X and Y be convex sets. Show that $Z = X \cap Y$ is also convex. What implication does this have for constructing constraint sets?

4 Show that the constraint sets discussed in this section (specifically, Eqs. (28), (29), (31), (34), (35), (36), (37)) are convex sets.

4 A Geometric Perspective: Convex Analysis

In this section, we focus on the view of the constraint set X as a geometric object embedded in $\mathbb{R}^n$. As an initial result, we have the following theorem on the existence of optimal minimizers:

Theorem 1 (The Extreme Value Theorem) *A continuous function $f : \mathbb{R}^n \mapsto \mathbb{R}$ over a closed, bounded set $X \subset \mathbb{R}^n$ attains its minimum, i.e., there is an $\underline{x}^* \in X$ such that $f(\underline{x}^*) = \min_{\underline{x} \in X} f(\underline{x})$.*

The proof of this is somewhat technical, and discussion of it is given in Appendix A.

The importance of this result can be illustrated by example. For instance, considering the function $f(x) = x^2$ over the interval $X = [-2,3]$, since the interval is closed and bounded, the theorem guarantees that a minimizer exists, and indeed we have $x^* = 0$. Consider however the function $f(x) = 1/(1+x)$ over the interval $X = [0,\infty)$. In this case, while the function f has a lower bound, i.e, $f(x) > 0$ for all $x \in X$, there is no $x^* \in X$ for which $f(x^*) = 0$ - no minimizer exists, and the conditions of the theorem are not met because the interval $[0,\infty)$ is not bounded. Consider as well the function $f(x) = x^2$ over the interval $X = (3,4)$. In this case again we have a lower bound, namely that $f(x) > 9$ for all $x \in X$, but there is no minimizer $x^* \in X$ such that $f(x^*) = 9$ - the conditions of the theorem are not met, because the interval $(3,4)$ is not closed.

This theorem guarantees the existence of a minimizer in most cases of interest, a result unavailable in the previous, unconstrained case, as continuous functions over $\mathbb{R}^n$ may or may not attain their minimizers. This result is therefore reassuring, but somewhat uninformative as to the nature of these minimizers. By assuming additional structure on f and X, we may derive more information about the optima.

In particular, we begin by taking $X \subset \mathbb{R}^n$ to be a closed, convex set. The following theorem gives a necessary condition any constrained minimum of f must satisfy:

Theorem 2 (Necessary Conditions) *If $\underline{x}^* \in X$ is a local minimum of f over X, and X is a closed, convex set, then for all $\underline{x}' \in X$,*

$$\nabla f(\underline{x}^*)^T(\underline{x}' - \underline{x}) \geqslant 0. \tag{39}$$

Proof. It is convenient to introduce the concept of a *feasible direction*:

Definition 2 *For a point $\underline{x} \in X$, a **feasible direction** $\underline{d}$ is one in which $\underline{x} + \alpha\underline{d} \in X$ for all sufficiently small values of $\alpha > 0$, i.e., a line may be extended from $\underline{x}$ in the direction $\underline{d}$, and remains in the set X for at least some interval.*

If $\underline{x} \in X$ is a local minimum of f, it must be that the directional derivative in any feasible direction from $\underline{x}$ must be non-negative, i.e., if $\underline{d}$ is a feasible direction from $\underline{x}$,

$$\nabla f(\underline{x})^{\mathrm{T}}\underline{d} \geqslant 0. \tag{40}$$

Observe then that for $\underline{x} \in X$, for any $\underline{x}' \in X$, $\underline{d} = \underline{x}' - \underline{x}$ is a feasible direction from $\underline{x}$. To see this, note that from the convexity of X we have that $\underline{x} + \alpha\underline{d} = (1-\alpha)\underline{x} + \alpha\underline{x}' \in X$ for any $0 \leqslant \alpha \leqslant 1$. The result follows immediately. □

This result can be viewed as a restricted version of the parallel result for unconstrained optimization, that at a minimum $\underline{x}^*$ over $\mathbb{R}^n$, $\nabla f(\underline{x}^*) = 0$. Note, if $\underline{x}^* \in X$ and any direction is a feasible direction, i.e., $\underline{x}^*$ is in the *interior* of X, it follows from the above theorem that $\nabla f(\underline{x}^*) = 0$ (consider the result for a feasible direction $\underline{d}$ and feasible direction $-\underline{d}$). Hence, the restricted result as stated in the above theorem is really only meaningful for local minima that lie on the boundary of $\underline{X}$.

An Application: The above necessary condition can be used to derive properties of a given minimizer. For instance, consider the minimizing f over the positive orthant

$$X = \{\underline{x} \in \mathbb{R}^n : x_i \geqslant 0 \text{ for } i = 1, 2, \ldots, n\}. \tag{41}$$

If $\underline{x}^* \in X$ is a minimizer, consider the hypothetical point $\underline{x}'$ such that $\underline{x}' = (x_1^*, x_2^*, \ldots, x_i^* + 1, \ldots, x_n^*)$ for a given i, that is $\underline{x}'$ agrees with $\underline{x}^*$ everywhere but in the i-th coordinate. Clearly, $\underline{x}' \in X$. In that case, applying the necessary condition above

$$\nabla f(\underline{x}^*)^{\mathrm{T}}(\underline{x}' - \underline{x}^*) = \frac{\partial f}{\partial x_i}(\underline{x}^*) \geqslant 0. \tag{42}$$

As the above must hold for all i, we have that the partial derivative with respect to each coordinate at the minimum must be non-negative.

Further, suppose for a given i we have that $x_i^* > 0$. Consider the hypothetical point $\underline{x}' = (x_1^*, \ldots, x_i^*/2, \ldots, x_n^*)$, that is $\underline{x}'$ agrees with $\underline{x}^*$ everywhere but the i-th coordinate, which is halved. In this case, applying the necessary condition, we have

$$\nabla f(\underline{x}^*)^{\mathrm{T}}(\underline{x}' - \underline{x}^*) = -\frac{1}{2} x_i^* \frac{\partial f}{\partial x_i}(\underline{x}^*) \geqslant 0, \tag{43}$$

or, since $x_i^* > 0$, we have that $\partial f / \partial x_i(\underline{x}) \leqslant 0$.

Combining the two results, we have that at any minimizer $\underline{x}^*$, $\partial f / \partial x_i(\underline{x}^*) = 0$ for any i such that $x_i^* > 0$, and $\partial f / \partial x_i(\underline{x}^*) \geqslant 0$ for all i.

If, as in the unconstrained case, place more structural restrictions on f, namely that f be a convex function over X, we may strengthen these results further:

Theorem 3 *If f is convex over X, then any local minimum over X is also a global minimum over X. Additionally, if f is **strictly** convex, then there is a unique global minimizer over X.*

Proof. The proof of this theorem goes exactly as in the unconstrained case, as presented in the Unconstrained Optimization Notes. □

4.1 Projections

A useful tool in the study of minimization from a geometric perspective is that of *projection*.

Definition 3 *Given a closed, convex set $X \subset \mathbb{R}^n$ and a point $\underline{z} \in \mathbb{R}^n$, the projection of $\underline{z}$ into X is defined to be the point in X that minimizes the distance to $\underline{z}$. That is, the optimal solution to*

$$\begin{aligned} &\textit{minimize } ||\underline{x} - \underline{z}||^2 \\ &\textit{s.t. } \underline{x} \in X. \end{aligned} \tag{44}$$

The projection of $\underline{z}$ into X will be denoted $[\underline{z}]^+$.

The projection has a number of important properties, summarized below.

Theorem 4 *Taking X as a closed, convex subset of $\mathbb{R}^n$, any $\underline{z} \in \mathbb{R}^n$, the following are true:*

i) The solution of the problem in Eq. (76) *exists and is **unique**, i.e., the projection $[\underline{z}]^+$ is well defined.*

ii) The projection $[\underline{z}]^+$ uniquely satisfies

$$(\underline{z} - [\underline{z}]^+)^T (\underline{x} - [\underline{z}]^+) \leqslant 0 \textit{ for all } \underline{x} \in X. \tag{45}$$

iii) For any $\underline{z}, \underline{w} \in \mathbb{R}^n$,

$$||[\underline{z}]^+ - [\underline{w}]^+|| \leqslant ||\underline{z} - \underline{w}||. \tag{46}$$

iv) If X is a subspace of $\mathbb{R}^n$, then $[\underline{z}]^+$ uniquely satisfies

$$\underline{x}^T(\underline{z} - [\underline{z}]^+) = 0 \text{ for all } x \in X, \tag{47}$$

i.e., $\underline{z} - [\underline{z}]^+$ is orthogonal to X.

Proof. The proof of (i), (ii), and (iv) are given as example problems below. The proof of each rely on the previous results on the minimizers of f over convex sets, in this case taking $f(\underline{x}) = ||\underline{x} - \underline{z}||^2 = (\underline{x} - \underline{z})^{\mathrm{T}}(\underline{x} - \underline{z})$. To prove (iii), we may utilize (ii) in the following way. Fixing $\underline{z}, \underline{w} \in \mathbb{R}^n$, we have that

$$(\underline{z} - [\underline{z}]^+)^{\mathrm{T}}([\underline{w}]^+ - [\underline{z}]^+) \leqslant 0, \tag{48}$$

since $[\underline{w}]^+ \in X$, and that

$$(\underline{w} - [\underline{w}]^+)^{\mathrm{T}}([\underline{z}]^+ - [\underline{w}]^+) \leqslant 0, \tag{49}$$

since $[\underline{z}]^+ \in X$.

Adding these inequalities yields

$$(\underline{z} - [\underline{z}]^+ - \underline{w} + [\underline{w}]^+)^{\mathrm{T}}([\underline{w}]^+ - [\underline{z}]^+) \leqslant 0, \tag{50}$$

or

$$(\underline{z} - \underline{w})^{\mathrm{T}}([\underline{w}]^+ - [\underline{z}]^+) - ([\underline{z}]^+ - [\underline{w}]^+)^{\mathrm{T}}([\underline{w}]^+ - [\underline{z}]^+) \leqslant 0, \tag{51}$$

which may be simplified to yield

$$||[\underline{z}]^+ - [\underline{w}]^+||^2 \leqslant (\underline{z} - w)^{\mathrm{T}}([\underline{z}]^+ - [\underline{w}]^+). \tag{52}$$

However, we have the result that for any vectors $\underline{u}, \underline{v} \in \mathbb{R}^n$, $\underline{u}^T \underline{v} \leqslant ||\underline{u}||||\underline{v}||$, which applied to the above yields

$$||[\underline{z}]^+ - [\underline{w}]^+||^2 \leqslant ||\underline{z} - \underline{w}||||[\underline{z}]^+ - [\underline{w}]^+||. \tag{53}$$

The result follows immediately from this. $\square$

Computing projections is of course a constrained minimization problem in itself. However, geometric intuition often renders it fairly straightforward in practice, especially for structurally simple X, such as spheres or boxes. For instance, taking X as the positive orthant, $X = \{\underline{x} \in \mathbb{R}^n : \forall i, x_i \geqslant 0\}$, the projection of any $\underline{z} \in \mathbb{R}$ is given by

$$[\underline{z}]^+ = (z_1^+, z_2^+, \ldots, z_n^+), \tag{54}$$

where z_i^+ represents the 'positive part' of z_i, i.e., gives z_i if $z_i \geqslant 0$, and 0 if $z_i < 0$.

Geometry frequently renders projection a useful and effective tool in practice.

4.2 Example Problems

1 Under the heading of unconstrained optimization, we have a theorem that says that if f is a *strictly* convex function, that any minimum is in fact the unique global minimum. Show that $f(x) = e^x$ is a strictly convex function, but that it has no minimizer over $X = \mathbb{R}$. How does this agree or disagree with the theorems presented in this section?

2 Consider applying the Necessary Condition theorem to minimizing a function f over a the simplex

$$X = \{\underline{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1 \text{ and } x_i \geqslant 0 \text{ for } i = 1, 2, \ldots, n\}. \tag{55}$$

What may be concluded? What are useful alternative points $\underline{x}' \in X$ to consider?

3 Prove items (i), (ii) and (iv) of Theorem 4. *Hint: For (i), to prove uniqueness, proceed by either by proving the strict convexity of the objective function in Eq.* (76) *or proceed by contradiction, assuming two distinct elements $\underline{x}_1, \underline{x}_2 \in X$ that minimize the distance to $\underline{z}$. Consider feasible directions at these points. For (ii), consider the necessary conditions for a minimizer to Eq.* (76). *For (iv), use item (ii) and the assumption that X is a subspace.*

4 Derive and prove a simple formula for the projection of a point $\underline{z}$ onto the sphere of radius R centered at $\underline{x}_0$,

$$X = \{\underline{x} \in \mathbb{R}^n : ||\underline{x} - \underline{x}_0|| \leqslant R\}. \tag{56}$$

5 Suppose that $X = \{\underline{x} \in \mathbb{R}^n : g(\underline{x}) \leqslant 0\}$ for a convex function g. Prove that for any $\underline{z} \notin X$, the vector $\underline{z} - [\underline{z}]^+$ is orthogonal to the plane tangent to X at $[\underline{z}]^+$.

6 Prove Eq. (54).

5 Descent Algorithm Analogs

In the case of unconstrained optimization, we considered descent algorithms that constructed sequences of points that converged to a minimum of f. These sequences were of the following form, iterating

$$\underline{x}_{k+1} = \underline{x}_k + \alpha_k \underline{d}_k \tag{57}$$

where $\underline{d}_k$ was a direction of descent chosen so that $\nabla f(\underline{x}_k)^{\mathrm{T}} \underline{d}_k < 0$ and $\alpha_k > 0$ was a propitiously chosen stepsize such that $f(\underline{x}_k + \alpha_k \underline{d}_k) < f(\underline{x}_k)$.

Applying similar algorithms in the constrained optimization case runs into the following difficulty, that the iterates must be constrained to lie in the feasible set X. Hence, directions and stepsizes must be chosen in such a way that the iterates remain feasible. In general, it suffices to ensure that $\underline{d}_k$ is a *feasible direction* at $\underline{x}_k$, as discussed previously. Under the assumption that X is convex, any feasible direction $\underline{d}$ at $\underline{x}$ may be expressed as $\underline{d} = \underline{x}' - \underline{x}$ for some $\underline{x}' \in X$. This allows the following generalization of the descent algorithm in the constrained case when applied to convex X: for a given $\underline{x}_k$, identify an $\underline{x}'_k \in X$ such that $\underline{x}'_k - \underline{x}_k$ is a descent direction. As the line between these two points is contained entirely within X, search along this line for a step that achieves descent. This yields a descent step

$$\underline{x}_{k+1} = \underline{x}_k + \alpha_k (\underline{x}'_k - \underline{x}_k), \tag{58}$$

where $\alpha_k \in (0, 1]$ is chosen to guarantee descent. It is possible that $\alpha_k > 1$ might achieve improved descent, but because of convexity we know that all stepsizes from 0 to 1 will yield a feasible $\underline{x}_{k+1} \in X$.

The problem of constructing a descent algorithm therefore becomes one of systematically generating appropriate $\underline{x}_k'$ and stepsizes α_k. For a given $\underline{x}_k'$, the stepsize α_k may be chosen a number of ways, but in particular it is worth mentioning:

- **Armijo's Rule:** This is exactly as stated in the unconstrained case; starting with an initial stepsize guess, the guess is reduced by some factor $0 < \beta < 1$ successively until a descent-achieving stepsize is found. In this case, we may take $s = 1$ as the initial guess at the stepsize.
- **Limited Minimization Rule:** Consider defining the function $g(s) = f(\underline{x}_k + s(\underline{x}_k' - \underline{x}_k))$. We may use any of the minimization rules discussed previously (for instance, quadratic interpolation) to identify the optimal s^* such that $g(s^*) = \min_{0 \leqslant s \leqslant 1} g(s)$, and then take $\alpha_k = s^*$.
- **Constant Stepsize:** Taking a constant stepsize of $\alpha_k = \alpha$ can frequently lead to convergence, however if α is chosen to be too small, convergence may be particularly slow, and if α is chosen to be too large, convergence cannot be guaranteed. This parallels the case in unconstrained optimization.

Note that as long as $\alpha_k \in (0,1]$, we are guaranteed that $\underline{x}_{k+1}$ will be feasible; the only concern is achieving descent over this range. We delay the discussion of systematically generating appropriate $\underline{x}_k'$ momentarily.

With regards to convergence, we may demonstrate similar guarantees as in the unconstrained case: namely if the sequence of feasible directions $\underline{d}_k = \underline{x}_k' - \underline{x}_k$ is **gradient related**, and the stepsizes are chosen by either of the above rules, any limit point of the resulting sequence $\{\underline{x}_k\}$ will satisfy the necessary conditions for a minimum as in Theorem 2. For a discussion of gradient related sequences, see the relevant section in the Unconstrained Optimization Notes.

We consider primarily two methods for generating useful $\underline{x}_k'$, the **conditional gradient method**, and the **gradient projection method**.

5.1 The Conditional Gradient Method

In this case, for a given $\underline{x}_k$, we consider $\underline{x}_k'$ to be the point in X that essentially maximizes both the rate of descent from $\underline{x}_k$, and the length of the interval over which to search (thus increasing the likelihood of finding a good descent step). In particular, we take $\underline{x}_k'$ as the solution to the following minimization problem

$$\begin{aligned} &\text{minimize } \nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}' - \underline{x}_k) \\ &\text{s.t. } \underline{x}' \in X. \end{aligned} \tag{59}$$

This represents a secondary optimization sub-problem that will need to be solved in every step of the algorithm. Fortunately, the above problem has some advantages which may make it relatively simple: for instance, note that the objective function is in fact linear in terms of the variables being optimized over, taking $\underline{x}_k$ as fixed. In this case, it may be simple to solve over various X; in the case of X defined by linear inequalities, this is simply a Linear Programming problem, for which many efficient solution algorithms exist. We observe that in many cases, the optimal solution will be an $\underline{x}'$ in a direction from $\underline{x}_k$ opposite the gradient $\nabla f(\underline{x}_k)$, as far away from $\underline{x}_k$ as allowed by X. This attempts to minimize the dot product of the gradient $\nabla f(\underline{x}_k)$ with the difference $\underline{x}' - \underline{x}_k$, but other superior solutions may be allowed by the geometry of X.

In general the convergence of this method can be quite poor, the errors converging to 0 at a sub-linear or sub-geometric rate, but is improved (to linear) when the constraint set X has 'sufficient curvature'.

Termination Condition: If f is convex, the following can be shown:

$$f(\underline{x}_k) \geqslant \min_{\underline{x}\in X} f(\underline{x}) \geqslant f(\underline{x}_k) + \nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}'_k - \underline{x}_k), \tag{60}$$

and that $\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}'_k - \underline{x}_k) \to 0$ as $k \to \infty$. Hence, $|\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}'_k - \underline{x}_k)| < \varepsilon$ may be used as a termination condition for this algorithm in this case.

5.2 The Gradient Projection Method

In this case, for some secondary stepsize s_k, we define

$$\underline{x}'_k = [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+, \tag{61}$$

and then take $\underline{x}_{k+1} = \underline{x}_k + \alpha_k(\underline{x}'_k - \underline{x}_k)$.

This is frequently implemented taking $\alpha_k = 1$, in which case we have the iteration:

$$\underline{x}_{k+1} = [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+. \tag{62}$$

This has a very nice, natural interpretation: we take a step of steepest descent with a given stepsize s_k, and then project back into the set X. In the case that the steepest descent step remains in X, the projection does nothing and we have simply taken a step of steepest descent. One of the surprising things about this algorithm is that $\underline{x}_{k+1} - \underline{x}_k$ is a descent direction, for all stepsizes s_k, unless $\underline{x}_k$ satisfies the necessary conditions for minima in Theorem 2.

Theorem 5 *For $\underline{x}_{k+1}$ as defined above,*

$$\nabla f(\underline{x}_k)^T(\underline{x}_{k+1} - \underline{x}_k) < 0, \tag{63}$$

unless $\underline{x}_k = [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+$, in which case $\underline{x}_k$ satisfies the necessary conditions for a minimum as in Theorem 2.

Proof. We begin by assuming $\underline{x}_k \neq [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+$, i.e., $\underline{x}_{k+1} \neq \underline{x}_k$. We will deal with this case separately.

We have the following:

$$\begin{aligned}
\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}_{k+1} - \underline{x}_k) &= \nabla f(\underline{x}_k)^{\mathrm{T}}([\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+ - \underline{x}_k) \\
&= \frac{1}{s_k}\left(-s_k \nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}_k - [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+)\right) \\
&= \frac{1}{s_k}\left(((\underline{x}_k - s_k \nabla f(\underline{x}_k)) - \underline{x}_k)^{\mathrm{T}}(\underline{x}_k - [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+)\right)
\end{aligned} \tag{64}$$

Let $\underline{z}_k = \underline{x}_k - s_k \nabla f(\underline{x}_k)$. In which case

$$\begin{aligned}
\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}_{k+1} - \underline{x}_k) &= \frac{1}{s_k}(\underline{z}_k - \underline{x}_k)^{\mathrm{T}}(\underline{x}_k - [\underline{z}_k]^+) \\
&= \frac{1}{s_k}((\underline{z}_k - [\underline{z}_k]^+) - (\underline{x}_k - [\underline{z}_k]^+))^{\mathrm{T}}(\underline{x}_k - [\underline{z}_k]^+) \\
&= \frac{1}{s_k}(\underline{z}_k - [\underline{z}_k]^+)^{\mathrm{T}}(\underline{x}_k - [\underline{z}_k]^+) - \frac{1}{s_k}||\underline{x}_k - [\underline{z}_k]^+||^2
\end{aligned} \tag{65}$$

Note, in the above we have that $(\underline{z}_k - [\underline{z}_k]^+)^{\mathrm{T}}(\underline{x}_k - [\underline{z}_k]^+) \leqslant 0$ by the properties of the projection (since $\underline{x}_k \in X$), and that $||\underline{x}_k - [\underline{z}_k]^+||^2 > 0$ by the initial assumption that $\underline{x}_k \neq \underline{x}_{k+1}$. Hence in this case, we have

$$\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}_{k+1} - \underline{x}_k) < 0. \tag{66}$$

Now consider the case that $\underline{x}_k = [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+$. We wish to verify the condition given in Theorem 2, i.e, $\underline{x}_k$ satisfies the necessary conditions for a minimum. Let $\underline{x} \in X$.

$$\begin{aligned}
\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x} - \underline{x}_k) &= \nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x} - [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+) \\
&= -\frac{1}{s_k}\left(-s_k \nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x} - [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+)\right) \\
&= -\frac{1}{s_k}\left(((\underline{x}_k - s_k \nabla f(\underline{x}_k)) - \underline{x}_k)^{\mathrm{T}}(\underline{x} - [\underline{x}_k - s_k \nabla f(\underline{x}_k)]^+)\right) \\
&= -\frac{1}{s_k}\left((\underline{z}_k - \underline{x}_k)^{\mathrm{T}}(\underline{x} - [\underline{z}_k]^+)\right) \\
&= -\frac{1}{s_k}\left((\underline{z}_k - [\underline{z}_k]^+)^{\mathrm{T}}(\underline{x} - [\underline{z}_k]^+)\right).
\end{aligned} \tag{67}$$

it follows from the above and the properties of the projection that $\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x} - \underline{x}_k)$, and that this holds for all $\underline{x} \in X$. □

To close, we consider the problem of the rate of convergence of the gradient projection method. The following is an important property of any minimizer $\underline{x}^* \in X$:

Proposition 1 *If $\underline{x}^*$ is a local minimum f over X, then for all $s > 0$,*

$$\underline{x}^* = [\underline{x}^* - s\nabla f(\underline{x}^*)]^+. \tag{68}$$

The proof is left as an exercise. We utilize this result in the following way.

We demonstrate that in general, we expect the gradient projection method to have linear or geometric convergence in error. Consider the case of $f(\underline{x}) = (1/2)\underline{x}^{\mathrm{T}} Q \underline{x} - \underline{b}^{\mathrm{T}} \underline{x}$ over a closed convex set X, with Q as a positive definite matrix with maximal and minimal eigenvalues given by $\lambda_{\max}, \lambda_{\min}$ respectively. Recall that in general, we expect any arbitrary function to behave like a quadratic form (defined by the Hessian) near a minimum, so this is a useful test case. Consider the gradient projection method with a constant stepsize s, converging to a minimum $\underline{x}^* \in X$. Then, utilizing the property of projection that it is non-expansive (see the relevant projection theorem),

$$\begin{aligned} ||\underline{x}_{k+1}-\underline{x}^*|| &= ||[\underline{x}_k - s\nabla f(\underline{x}_k)]^+ - [\underline{x}^* - s\nabla f(\underline{x}^*)]^+|| \\ &\leqslant ||(\underline{x}_k - s\nabla f(\underline{x}_k)) - (\underline{x}^* - s\nabla f(\underline{x}^*))|| \\ &= ||\underline{x}_k - s\nabla f(\underline{x}_k)) - \underline{x}^* + s\nabla f(\underline{x}^*)|| \\ &= ||\underline{x}_k - \underline{x}^* - s(Q\underline{x}_k - Q\underline{x}^*)|| \\ &= ||(\mathrm{I}-sQ)(\underline{x}_k - \underline{x}^*)|| \\ &\leqslant \max(|1-s\lambda_{\max}|,|1-s\lambda_{\min}|)||\underline{x}_k - \underline{x}^*||. \end{aligned} \tag{69}$$

The last step utilizes the fact that if λ is an eigenvalue of Q, $1-s\lambda$ is an eigenvalue of $\mathrm{I}-sQ$, and an eigenvalue bound discussed in the Unconstrained Optimization Notes.

Hence from the above, we see that if s is sufficiently small, the error on $\underline{x}_{k+1}$ is at most some constant factor less than 1 of the error on $\underline{x}_k$. This guarantees a worst case linear or geometric convergence of error, if s is sufficiently small.

5.2.1 Alternative Stepsize Choices

The above convergence analysis was predicated on taking a constant stepsize $s_k = s$. Naturally this is not the only choice. However, consider attempting to find the absolute optimal stepsize, i.e., trying to find

$$s_k = \arg\min_{s>0} f([\underline{x}_k - s\nabla f(\underline{x}_k)]^+). \tag{70}$$

While the function $f([\underline{x}_k - s\nabla f(\underline{x}_k)]^+)$ can be shown to be continuous, it will likely lack differentiability at all points, or possess other analytic quirks that make successive approximation (and therefore estimation of an optimal s^* difficult. As a systematic alternative, we might consider adapting Armijo's Rule, giving *Armijo's Rule Along the Projection Arc*:

For an initial stepsize guess $s' > 0$, contraction factor $\beta \in (0,1)$ and cutoff $\sigma \in (0,1)$, consider finding the smallest integer $t \geqslant 0$ such that

$$\frac{f(\underline{x}_k) - f([\underline{x}_k - s'\beta^t\nabla f(\underline{x}_k)]^+)}{\nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}_k - [\underline{x}_k - s'\beta^t\nabla f(\underline{x}_k)]^+)} \geqslant \sigma, \tag{71}$$

and then take $s_k = s'\beta^t$.

As with the classical Armijo's Rule, this essentially starts with an initial guess at a good stepsize, and reduces it by a constant factor until sufficient descent is achieved. The only computational burden (aside from the successive evaluations of f) is in computing the projection $[\underline{x}_k - s'\beta^t\nabla f(\underline{x}_k)]^+$.

5.3 Additional Methods

One additional method worth considering is the constrained analog of Newton's method. In this case, we have

$$\underline{x}_{k+1} = \underline{x}_k + \alpha_k(\underline{x}'_k - \underline{x}_k), \tag{72}$$

and

$$\underline{x}_k' = \arg\min_{\underline{x}\in X} \left[f(\underline{x}_k) + \nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}-\underline{x}_k) + \frac{1}{2s_k}(\underline{x}-\underline{x}_k)^{\mathrm{T}}\nabla^2 f(\underline{x}_k)(\underline{x}-\underline{x}_k) \right]. \tag{73}$$

In particular, taking $\alpha_k = 1, s_k = 1$, we have

$$\underline{x}_{k+1} = \arg\min_{\underline{x}\in X} \left[f(\underline{x}_k) + \nabla f(\underline{x}_k)^{\mathrm{T}}(\underline{x}-\underline{x}_k) + \frac{1}{2}(\underline{x}-\underline{x}_k)^{\mathrm{T}}\nabla^2 f(\underline{x}_k)(\underline{x}-\underline{x}_k) \right], \tag{74}$$

which is to say, we take $\underline{x}_{k+1}$ to be the point in X that minimizes the second order Taylor approximation of f, centered at $\underline{x}_k$. This is the natural 'constrained' extension of the classical Newton's method. As such, if it is started with some $\underline{x}_0$ sufficiently close to a minimum, convergence should be (and is generally) quick and effective - indeed, under some assumptions on f, convergence can again be shown to be superlinear. If it is unclear where a good general starting point may be, considering the more general framework with flexible stepsizes α_k, s_k (potentially chosen by Armijo's Rules, as appropriate) may be more effective.

However, computationally speaking, computing the solution to the above constrained minimization problem (while quadratic forms as above are generally as simple as one might hope for) can prove difficult, particularly for complex X but even for some simple X.

5.4 Example Problems

1. Revisit the proof of convergence for descent algorithms in the unconstrained case. Generalize it to the constrained case as above.
2. Let X be the unit sphere centered at the origin. For a given $\underline{x} \in X$, find the largest $s \geqslant 0$ such that $\underline{x} - s\nabla f(\underline{x}) \in X$.
3. Prove Eq. (60).
4. Implement the gradient projection method for a suitable test f over the unit sphere, for various choices of stepsize or stepsize rule for s_k. What is a good initial guess? How does it perform? What effect does the choice of stepsize or stepsize rule have?
5. What are good termination conditions for the gradient projection method?
6. Prove Proposition 1.
7. As an example of the potential difficulties implementing a constrained Newton's method, consider the problem of minimizing the quadratic form $(1/2)(\underline{x}-\underline{x}_0)^T Q(\underline{x}-\underline{x}_0)$ for a symmetric, positive definite matrix Q and some point $\underline{x}_0$, where the constraint set X is defined to be the unit ball centered at the origin. Assume that $\underline{x}_0 \notin X$.

6 An Analytic Perspective: Lagrange Multipliers

In this section, we view the feasible set $X \subset \mathbb{R}^n$ from an analytic perspective, as defined by a system of equalities and inequalities. In particular, we take a set of functions $h_i : \mathbb{R}^n \mapsto \mathbb{R}$ for $i = 1, \ldots, m$

and $g_j : \mathbb{R}^n \mapsto \mathbb{R}$ for $j = 1, \dots, r$, and consider sets of the form

$$X = \{\underline{x} \in \mathbb{R}^n : \forall i, h_i(\underline{x}) = 0 \text{ and } \forall j, g_j(\underline{x}) \leqslant 0\}. \tag{75}$$

As noted previously, it is possible to exchange between equality and inequality constraints: $h_i(\underline{x}) = 0$ becomes $h_i(\underline{x}) \leqslant 0$ and $-h_i(\underline{x}) \leqslant 0$; $g_j(\underline{x}) \leqslant 0$ becomes, through the addition of new dummy or slack variables, $g_j(\underline{x}) + (x'_j)^2 = 0$. Some computational approaches will be more applicable to one form or another, so it is good to have a sense of how to translate between them.

It is convenient to define the vector functions $\underline{h}(\underline{x}) = (h_1(\underline{x}), \dots, h_m(\underline{x}))$ and $\underline{g}(\underline{x})$ defined similarly. Hence the constraints may be rewritten as $\underline{h}(\underline{x}) = 0$ and $\underline{g}(\underline{x}) \leqslant 0$.

In the unconstrained case, we derived the condition that at any minimum $\underline{x}^* \in \mathbb{R}^n$, we must have $\nabla f(\underline{x}^*) = 0$ (and related conditions on the Hessian). This was our primary analytic tool for studying minima - many of the algorithms discussed can be thought of in terms of attempting to solve the system of equations defined by $\nabla f(\underline{x}) = 0$. In this section, we derive a similar condition for minima that similarly gives a system of equations that can be solved, analytically or algorithmically, to yield minima.

6.1 Equality Constraints

In this section, we consider the case of equality constraints alone, that is:

$$\begin{aligned} \text{minimize } & f(\underline{x}) \\ \text{s.t. } & \underline{h}(\underline{x}) = 0. \end{aligned} \tag{76}$$

We consider inequality constraints in the next subsection.

The main result here is the classical Lagrange multipliers theorem. First, we define the following concept:

Definition 4 *A point $\underline{x} \in \mathbb{R}^n$ is* ***regular*** *if the vectors $\{\nabla h_1(\underline{x}), \nabla h_2(\underline{x}), \dots, \nabla h_m(\underline{x})\}$ are linearly independent.*

We have the following theorem:

Theorem 6 (Lagrange Multiplier Necessary Conditions for Minima) *Let $\underline{x}^*$ be a local minimum of f subject to $\underline{h}(\underline{x}^*) = 0$, and let $\underline{x}^*$ be regular. Then there exists a unique vector of constants $\underline{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*) \in \mathbb{R}^m$ such that*

$$\nabla f(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(\underline{x}^*) = 0. \tag{77}$$

Additionally, if f and $\underline{h}$ are twice continuously differentiable,

$$\underline{d}^T \left(\nabla^2 f(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(\underline{x}^*) \right) \underline{d} \geqslant 0, \tag{78}$$

for all vectors $\underline{d} \in \mathbb{R}^n$ such that $\underline{d}^T \nabla h_i(\underline{x}^*) = 0$ for all $i = 1, \dots m$.

In short, this result provides a system of equations of the form

$$\left\{ \frac{\partial f}{\partial x_k}(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i^* \frac{\partial h_i}{\partial x_k}(\underline{x}^*) = 0 : \; k = 1, 2, \ldots, n \right\} \tag{79}$$

and

$$\{h_i(\underline{x}^*) = 0 : \; i = 1, \ldots, m\} \tag{80}$$

that may be solved simultaneously ($n + m$ equations in n unknown $\underline{x}$-coordinates and m unknown $\underline{\lambda}$-coordinates) to yield potential minima of the function, the same way $\nabla f(\underline{x}) = 0$ was solved in the unconstrained case. The second order condition on the Hessians provides a mechanism for distinguishing maxima from minima, as the $\nabla^2 f(\underline{x}^*)$ as positive semi-definite condition did, in the unconstrained case.

There is a lot to unpack about this theorem, but it is useful to see some examples of its use before attempting to explain or prove it.

6.1.1 Examples for Lagrange Multipliers with Equality Constraints

Note: in each of the following examples, the 'regularity' condition of the minimizer in the Lagrange multipliers theorem is immediately satisfied, since there is only one equality constraint in each case.

- Consider the following problem, for $R > 0$:

$$\begin{aligned} &\text{minimize } x_1 + x_2 + \ldots + x_n \\ &\text{s.t. } \sum_{i=1}^{n} x_i^2 = R^2. \end{aligned} \tag{81}$$

  This may be expressed taking $f(\underline{x}) = x_1 + \ldots + x_n$ and $h(\underline{x}) = \underline{x}^{\mathrm{T}}\underline{x} - R^2$. Applying the Lagrange multiplier theorem, we have that any minimum $\underline{x}$ has some scalar value λ such that $h(\underline{x}) = 0$ and

$$\{1 + 2\lambda x_i = 0 : \; i = 1, \ldots n\}. \tag{82}$$

  From the Lagrange multiplier equations above, we have that $\lambda \neq 0$, and that for each i, $x_i = -1/(2\lambda)$. The constraint $h(\underline{x}) = 0$ therefore gives us $n(-1/(2\lambda))^2 = R^2$ or $-1/(2\lambda) = \pm R/\sqrt{n}$. This yields two feasible solutions for $\underline{x}$:

$$\begin{aligned} \underline{x}_1 &= (R/\sqrt{n}, \ldots, R/\sqrt{n}) : f(\underline{x}_1) = R\sqrt{n} \\ \underline{x}_2 &= (-R/\sqrt{n}, \ldots, -R/\sqrt{n}) : f(\underline{x}_2) = -R\sqrt{n}. \end{aligned} \tag{83}$$

  As the above represent the only feasible solutions to the Lagrange multiplier equations, and therefore the only feasible minima, we see from comparing the function values that $\underline{x}^* = \underline{x}_2$ is a global minimum.

  We could also have utilized the second order conditions in the following way: note that $\nabla^2 f(\underline{x}) = 0$ and $\nabla^2 h(\underline{x}) = 2\mathrm{I}$ for all $\underline{x}$, hence $\nabla^2 f(\underline{x}) + \lambda \nabla^2 h(\underline{x})$ is in fact positive semi definite if $\lambda \geqslant 0$, which yields only one feasible minimum: $\lambda^* = \sqrt{n}/(2R)$ and $\underline{x}^* = \underline{x}_2$ as above.

- Consider the following problem, for a given $\underline{x}_0, \underline{u} \in \mathbb{R}^n$:

$$\begin{aligned} \text{minimize } & x_1^2 + x_2^2 + \ldots + x_n^2 \\ \text{s.t. } & \underline{u}^{\mathrm{T}}(\underline{x} - \underline{x}_0) = 0. \end{aligned} \tag{84}$$

  Note, this can be interpreted as looking for the projection of the point $\underline{z} = 0$ onto the hyperplane that passes through $\underline{x}_0$ and is orthogonal to $\underline{u}$. Taking $f(\underline{x}) = \underline{x}^{\mathrm{T}}\underline{x}$ and $h(\underline{x}) = \underline{u}^{\mathrm{T}}(\underline{x} - \underline{x}_0)$, we have that any minimizer must satisfy $h(\underline{x}) = 0$ as well as

$$\nabla f(\underline{x}) + \lambda \nabla h(\underline{x}) = 0 \tag{85}$$

  or

$$2\underline{x} + \lambda \underline{u} = 0. \tag{86}$$

  The above gives us that $\underline{x} = -(\lambda/2)\underline{u}$, i.e., the minimizer must be in the direction (or opposite, depending on the sign of λ) of $\underline{u}$. The constraint $h(\underline{x}) = 0$ therefore gives us

$$\underline{u}^{\mathrm{T}}(-(\lambda/2)\underline{u} - \underline{x}_0) = 0 \tag{87}$$

  or

$$\lambda = -2\frac{\underline{u}^{\mathrm{T}}\underline{x}_0}{||u||^2} \tag{88}$$

  As the above is the unique solution for feasible λ, it must be the λ^* corresponding to the minimizer, hence we have

$$\underline{x}^* = \frac{\underline{u}^{\mathrm{T}}\underline{x}_0}{||u||^2}\underline{u}. \tag{89}$$

From the above cases, it can be extrapolated that when dealing with linear or quadratic or similar constraints and objective functions, the Lagrange multiplier equations are frequently algebraically solvable. In general, however, they can produce a system of non-linear equations that, along with the constraint equations themselves, can be difficult to solve. This prompts computational and algorithmic approaches, as it did in the case for unconstrained optimization.

6.1.2 Discussion and Proof of the Lagrange Multiplier Theorem

To motivate the Lagrange Multiplier result, consider the case in a single dimension, which is to say: at any local minimum $\underline{x}^*$ of f satisfying $h(\underline{x}^*) = 0$, $\nabla f(\underline{x}^*) + \lambda^* \nabla h(\underline{x}^*) = 0$ for some unique λ^* (the regularity condition is not a concern when there is a single constraint).

The equation $h(\underline{x}) = 0$ can be taken as defining a surface in $\mathbb{R}^n$ (as $\underline{x}^T\underline{x} - R^2$ defines a sphere of radius R). Any point on this surface is feasible for the problem. We may solving the problem in the following way: imagine a particle that is free to move over this feasible surface, and imagine that it is being 'dragged' by a current though space - when the particle is at $\underline{x}$, it feels a pull on it proportional to $-\nabla f(\underline{x})$. If there were no restriction h and the particle was free to move through space, this would drag it towards a local minimum of f, and it would stop moving there, when $\nabla f(\underline{x}^*) = 0$ and it feels no current.

But taking the particle as restricted to move only on the surface defined by h, how can this current pull it? At a differential level, if h is continuously differentiable, the only feasible directions of motion are in directions tangent to the surface - any other direction immediately will pull the particle off the surface, and is therefore not feasible (consider being restricted to move along a circle, for example). It can be shown that at every point on the surface, the gradient $\nabla h(\underline{x})$ is *normal* to the surface, which means that the feasible tangent directions are $\underline{d}$ such that $\nabla h(\underline{x})^{\mathrm{T}}\underline{d} = 0$. These directions define the plane tangent to h at $\underline{x}$.

With the particle restricted to move along the surface, if the particle is experiencing a pull proportional to $-\nabla f(\underline{x})$, and at least some of that force is in a feasible direction $\underline{d}$ (i.e., $(-\nabla f(\underline{x}))^{\mathrm{T}}\underline{d} > 0$), then the particle will move in that direction. The particle will stop moving at a point $\underline{x}^*$ when there is no more force in any feasible direction, i.e., $(-\nabla f(\underline{x}^*))^{\mathrm{T}}\underline{d} = 0$ for all feasible $\underline{d}$ at $\underline{x}^*$. Note, as the rate of change of f in any feasible direction is 0 (and since we arrived at the point following the negative gradient of f), we have that this stationary point $\underline{x}^*$ is a local minimum.

But, the condition $(-\nabla f(\underline{x}^*))^{\mathrm{T}}\underline{d} = 0$ for all $\underline{d}$ such that $\nabla h(\underline{x}^*)^{\mathrm{T}}\underline{d} = 0$, i.e., $-\nabla f(\underline{x}^*)$ has no projection in any direction orthogonal to $\nabla h(\underline{x}^*)$, implies that the vectors $\nabla f(\underline{x}^*)$ and $\nabla h(\underline{x}^*)$ must be parallel to each other. Another way of stating this is that one is a scalar multiple of the other, $-\nabla f(\underline{x}^*) = \lambda \nabla h(\underline{x}^*)$, or that for some unique λ^*,

$$\nabla f(\underline{x}^*) + \lambda^* \nabla h(\underline{x}^*) = 0. \tag{90}$$

To some extent, the model above can be generalized to being restricted to multiple surfaces $\{h_i(\underline{x}) = 0\}$, which the result of the theorem: the gradient of f at a minimum must lie in the span of the gradients of the constraints $\{\nabla h_i\}$, or equivalently, the gradient of f at a minimum $\underline{x}^*$ must be orthogonal to all directions of feasible motion at $\underline{x}^*$. Note, the assumption of regularity (linear independence of the $\nabla h_i(\underline{x}^*)$) ensures that the coefficients expressing $\nabla f(\underline{x}^*)$ in terms of the $\nabla h_i(\underline{x}^*)$ are unique.

Hopefully, this provides some intuition as to why the result holds, and renders it something less than magical. But the result may also be proven precisely. The full prove will not be presented in these notes, but the idea behind it is useful, as motivation for and preview of an algorithmic approach to follow. We return to the multiple constraint case, with the vector function $\underline{h}(\underline{x})$ containing all the individual $\{h_i\}$.

Essentially, consider defining a sequence of objective functions $\{F_k : \mathbb{R}^n \mapsto \mathbb{R}\}$ by

$$F_k(\underline{x}) = f(\underline{x}) + \frac{1}{2}k||\underline{h}(\underline{x})||^2 = f(\underline{x}) + \frac{1}{2}k\sum_{i=1}^{m} h_i(\underline{x})^2. \tag{91}$$

Consider the problem of minimizing F_k as a sequence of **unconstrained** optimization problems. Note, we are essentially augmenting the objective function f with a **penalty** for violating the constraints - as $k \to \infty$, the 'cost' of having large, or even non-zero values of the h_i grows. Hence, minima of the F_k will be increasingly forced towards points that both minimize f, and satisfy the constraints $\{h_i = 0\}$.

To simplify the presentation of the proof somewhat, for each k let $\underline{x}_k^*$ be an unconstrained local minimum of F_k, and assume that the sequence $\{\underline{x}_k^*\}$ converges, $\underline{x}_k^* \to \underline{x}^*$ where $\underline{x}^*$ is a regular, constrained local minimum of f satisfying $\underline{h}(\underline{x}^*) = 0$. (Note: the meat of the full proof is the construction of such a sequence $\{\underline{x}_k^*\}$ with those properties.)

Note then that we have for each k,

$$0 = \nabla F_k(\underline{x}_k^*) = \nabla f(\underline{x}_k^*) + k \nabla \underline{h}(\underline{x}_k^*) \underline{h}(\underline{x}_k^*), \tag{92}$$

where $\nabla \underline{h}(\underline{x}_k^*)$ is m-column matrix where column i is the gradient $\nabla h_i(\underline{x}_k^*)$.

Note, based on the assumption that $\underline{x}^*$ is regular, i.e., $\nabla h_1(\underline{x}^*), \ldots, \nabla h_m(\underline{x}^*)$ are linearly independent, since $\underline{x}_k^* \to \underline{x}^*$ we have that $\nabla h_1(\underline{x}_k^*), \ldots, \nabla h_m(\underline{x}_k^*)$ are linearly independent for sufficiently large k, hence the matrix $\nabla \underline{h}(\underline{x}_k^*)$ has linearly independent columns. By Prop. 4 in Section B, we have therefore that $\nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}_k^*)$ is positive definite, and therefore invertible.

Multiplying the above equation on the left by $[\nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}_k^*)]^{-1} \nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}}$, we have

$$\begin{aligned} 0 &= [\nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}_k^*)]^{-1} \nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla f(\underline{x}_k^*) + k [\nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}_k^*)]^{-1} \nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}_k^*) \underline{h}(\underline{x}_k^*) \\ &= [\nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}_k^*)]^{-1} \nabla \underline{h}(\underline{x}_k^*)^{\mathrm{T}} \nabla f(\underline{x}_k^*) + k\, \underline{h}(\underline{x}_k^*). \end{aligned} \tag{93}$$

From the above, we have the limit that as $k \to \infty$, (as $\underline{x}_k^* \to \underline{x}^*$ and f, $\underline{h}$ are continuous)

$$k\, \underline{h}(\underline{x}_k^*) \to -[\nabla \underline{h}(\underline{x}^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}^*)]^{-1} \nabla \underline{h}(\underline{x}^*)^{\mathrm{T}} \nabla f(\underline{x}^*). \tag{94}$$

Defining the vector $\underline{\lambda}^* = -[\nabla \underline{h}(\underline{x}^*)^{\mathrm{T}} \nabla \underline{h}(\underline{x}^*)]^{-1} \nabla \underline{h}(\underline{x}^*)^{\mathrm{T}} \nabla f(\underline{x}^*)$, and returning to

$$0 = \nabla F_k(\underline{x}_k^*) = \nabla f(\underline{x}_k^*) + k \nabla \underline{h}(\underline{x}_k^*) \underline{h}(\underline{x}_k^*), \tag{95}$$

we have in the limit as $k \to \infty$,

$$0 = \nabla f(\underline{x}^*) + \nabla \underline{h}(\underline{x}^*) \underline{\lambda}^*. \tag{96}$$

Note, based on the assumption that $\nabla h_1(\underline{x}^*), \ldots, \nabla h_m(\underline{x}^*)$ are linearly independent, the vector $\underline{\lambda}^*$ above is unique. The second order conditions (which distinguish a stationary point as a minimum rather than a maximum) involve slightly more work, but again the starting point is analysis of the unconstrained minima of the penalized objective function, which will have positive semi-definite Hessians of F_k.

The full proof is worth looking up, but the takeaway should primarily be the perspective of constrained minima as limits of the unconstrained minima. This immediately suggests an algorithmic approach for constrained optimization problems, which will be explored in the next section.

6.2 The Lagrangian Function and Sufficient Conditions

It is convenient to define the following function:

Definition 5 (The Lagrangian for Equality Constraints) *The **Lagrangian** of a constrained optimization problem for objective function f and equality constraints $\underline{h}(\underline{x}) = 0$ is given by*

$$L(\underline{x}, \underline{\lambda}) = f(\underline{x}) + \underline{\lambda}^T \underline{h}(\underline{x}). \tag{97}$$

Note that the Lagrangian is a linear function of the individual λ-coordinates, and $\nabla_\lambda L(\underline{x},\underline{\lambda}) = \underline{h}(\underline{x})$. This function allows for a concise restatement of the necessary conditions theorem:

Theorem 7 (Lagrange Multiplier Necessary Conditions for Minima) *Let $\underline{x}^*$ be a local minimum of f subject to $\underline{h}(\underline{x}^*) = 0$, and let $\underline{x}^*$ be regular. Then there exists a unique vector of constants $\underline{\lambda}^* = (\lambda_1^*, \ldots, \lambda_m^*) \in \mathbb{R}^m$ such that*

$$\begin{aligned} \nabla_x L(\underline{x}^*, \underline{\lambda}^*) &= 0 \\ \nabla_\lambda L(\underline{x}^*, \underline{\lambda}^*) &= 0. \end{aligned} \tag{98}$$

Additionally, if f and $\underline{h}$ are twice continuously differentiable,

$$\underline{d}^T \nabla_{xx}^2 L(\underline{x}^*, \underline{\lambda}^*) \underline{d} \geqslant 0 \tag{99}$$

for all $\underline{d} \in \mathbb{R}^n$ such that $\nabla \underline{h}(\underline{x}^*)^T \underline{d} = 0$.

Additionally, we state but do not prove the following result, which gives sufficient conditions for a minimizer. This result parallels the sufficient conditions in the unconstrained case, as the necessary conditions above parallel the necessary conditions in the unconstrained case.

Theorem 8 (Lagrange Multiplier Sufficient Conditions for Minima) *For f, $\underline{h}$ twice continuously differentiable, let $\underline{x}^* \in \mathbb{R}^n$ and $\underline{\lambda}^* \in \mathbb{R}^m$ satisfy*

$$\begin{aligned} \nabla_x L(\underline{x}^*, \underline{\lambda}^*) &= 0 \\ \nabla_\lambda L(\underline{x}^*, \underline{\lambda}^*) &= 0. \end{aligned} \tag{100}$$

and for all $\underline{d} \in \mathbb{R}^n$ such that $\underline{d} \neq 0$ and $\nabla \underline{h}(\underline{x}^*)^T \underline{d} = 0$, we have

$$\underline{d}^T \nabla_{xx}^2 L(\underline{x}^*, \underline{\lambda}^*) \underline{d} > 0. \tag{101}$$

In this case, $\underline{x}^*$ is a strict local minimum of f satisfying $\underline{h}(\underline{x}^*) = 0$.

6.3 Mixed Constraints

The results and analysis of the previous section extend fairly immediately to the case of constraints defined by both equalities and inequalities. Let $\underline{h}(\underline{x}) = 0$ represent the system of $h_i(\underline{x}) = 0$ constraints and $\underline{g}(\underline{x}) \leqslant 0$ represent the system of $g_j(\underline{x}) \leqslant 0$ constraints.

We define the following Lagrangian function for $\underline{x} \in \mathbb{R}^n, \underline{\lambda} \in \mathbb{R}^m, \underline{\mu} \in \mathbb{R}^r$:

$$\begin{aligned} L(\underline{x}, \underline{\lambda}, \underline{\mu}) &= f(\underline{x}) + \sum_{i=1}^{m} \lambda_i h_i(\underline{x}) + \sum_{j=1}^{r} \mu_j g_j(\underline{x}) \\ &= f(\underline{x}) + \underline{\lambda}^{\mathrm{T}} \underline{h}(\underline{x}) + \underline{\mu}^{\mathrm{T}} \underline{g}(\underline{x}). \end{aligned} \tag{102}$$

Additionally, it is convenient to define the concept of *active constraints*:

Definition 6 *For a system of inequalities $\underline{g}(\underline{x}) \leqslant 0$, at a given point $\underline{x} \in \mathbb{R}^n$, the* ***active constraints*** *are the indices $j = 1, \ldots, r$ which realize the corresponding inequality constraint with equality:*

$$A(\underline{x}) = \{j : g_j(\underline{x}) = 0\}. \tag{103}$$

Note that if $\underline{x}^*$ is a constrained local minimum and j is an *inactive* constraint, then $\underline{x}^*$ will also be a constrained local minimum of the problem where the constraint $g_j(\underline{x}) \leqslant 0$ is discarded. For any *active* constraint, however, discarding that constraint may result in a problem where $\underline{x}^*$ is not a constrained local minimum. As such, active inequality constraints function much like equality constraints, and inactive inequality constraints, roughly speaking, do not matter.

The importance of active versus inactive constraints will be discussed further shortly.

We have the following result:

Theorem 9 (The Karush-Kuhn-Tucker Optimality Conditions) *Let $\underline{x}^* \in \mathbb{R}^n$ be a constrained local minimum of f subject to $\underline{h}(\underline{x}^*) = 0$ and $\underline{g}(\underline{x}^*) \leqslant 0$.*

If $\underline{x}^*$ is ***regular*** *with respect to the $\{\nabla h_i(\underline{x}^*)\}$ as well as the gradients of the* ***active*** *inequality constraint gradients $\{\nabla g_j(\underline{x}^*)\}_{j \in A(\underline{x}^*)}$, then there is a unique pair of vectors $\underline{\lambda}^* \in \mathbb{R}^m$ and $\underline{\mu}^* \in \mathbb{R}^r$ such that*

$$\begin{aligned} \nabla_x L(\underline{x}^*, \underline{\lambda}^*, \underline{\mu}^*) &= 0 \\ \mu_j^* &\geqslant 0 \text{ for all } j = 1, \ldots r \\ \mu_j^* &= 0 \text{ for all } j \notin A(\underline{x}^*). \end{aligned} \tag{104}$$

or equivalently,

$$\begin{aligned} \nabla_x L(\underline{x}^*, \underline{\lambda}^*, \underline{\mu}^*) &= 0 \\ \mu_j^* g_j(\underline{x}^*) &= 0 \text{ for all } j = 1, \ldots r. \end{aligned} \tag{105}$$

Additionally, if $f, \underline{h}, \underline{g}$ are twice continuously differentiable, then

$$\underline{d}^T \nabla_{xx}^2 L(\underline{x}^*, \underline{\lambda}^*, \underline{\mu}^*) \underline{d} \geqslant 0 \tag{106}$$

for all $\underline{d} \in \mathbb{R}^n$ such that $\nabla \underline{h}(\underline{x}^*)^T \underline{d} = 0$ and $\nabla g_j(\underline{x}^*)^T \underline{d} = 0$ for all $j \in A(\underline{x}^*)$.

Note, there is a sufficient condition version of the above result, which parallels Theorem 8 as a sufficient condition version of Theorem 6, with the added assumption that the multipliers for active inequality constraints be positive.

The above result parallels fairly well the result for the equality constraint case, with a couple of exceptions worth noting: i) the Lagrange multipliers for *inactive* inequality constraints are specifically 0, and the inactive constraints do not enter into the second order conditions - this effectively corresponds to the earlier observation that inactive inequality constraints can effectively be discarded without altering a given solution to the problem; ii) the Lagrange multipliers for active inequality constraints are restricted to be non-negative.

The proof will not be presented here, but much of the discussion of the equality constraint case in Section 6.1.2 applies here. A proof of the above result can be given in much the same way, defining a penalized objective function that only penalizes $g_j(\underline{x}) > 0$.

In terms of the discussion in Section 6.1.2, however, the concept of 'feasible directions of motion as allowed by the constraints' justifies the non-negativity assumption on the inequality constraint multipliers: in particular, for a given constraint $h_i = 0$, the feasible directions of motion are those $\underline{d} \in \mathbb{R}^n$ that are in the tangent plane to the surface at $\underline{x}$, or in other words are orthogonal to the normal vector, $\nabla h_i(\underline{x})^{\mathrm{T}} \underline{d} = 0$. For active inequality constraints, however, the range of feasible motion is much larger: if $\underline{x}$ is on the surface $g_j = 0$, then any direction $\underline{d}$ that is in the tangent plane to the surface at $\underline{x}$ is feasible, *as well as any direction $\underline{d}$ below the tangent plane, into the volume defined by $g_j \leqslant 0$*. Note, for any $\underline{x}$ in the interior of $g_j \leqslant 0$, all directions are feasible directions.

This lends a significance to the orientation of the gradients $\{\nabla g_j(\underline{x})\}$, reflected in the non-negativity conditions on the Lagrange multipliers. This non-negativity will be addressed again in the future discussion on **sensitivity**.

Utilizing the KKT conditions to solve for minima can often develop a somewhat combinatorial flavor, in terms of looking at what different solutions arise assuming which constraints are active or inactive. Consider the following example: for non-trivial $\underline{u}, \underline{v} \in \mathbb{R}^n$, and $b_u, b_v \in \mathbb{R}$, with $b_u, b_v < 0$, and further let us take $\underline{u}$ and $\underline{v}$ to be linearly independent,

$$\begin{aligned} \text{minimize } & x_1^2 + x_2^2 + \ldots + x_n^2 \\ \text{s.t. } & \underline{u}^{\mathrm{T}} \underline{x} \leqslant b_u \\ & \underline{v}^{\mathrm{T}} \underline{x} \leqslant b_v. \end{aligned} \tag{107}$$

Each constraint above defines a half of $\mathbb{R}^n$ by defining a plane and restricting $\underline{x}$ to lie on one side of that plane. The full set of feasible $\underline{x}$ is therefore the intersection of these two half-spaces. Let $g_u(\underline{x}) = \underline{u}^{\mathrm{T}} \underline{x} - b_u$ and $g_v(\underline{x}) = \underline{v}^{\mathrm{T}} \underline{x} - b_v$.

We may proceed by cases, based on which constraints we take to be active:

$A(\underline{x}^*) = \emptyset$: In this case, taking both the u and v constraints as inactive, their Lagrange multipliers are taken to be 0, and the necessary condition becomes

$$\nabla f(\underline{x}) = 0 \tag{108}$$

which is solved uniquely by $\underline{x} = 0$. However, in this case $\underline{u}^{\mathrm{T}} \underline{x} - b_u = -b_u > 0$ and $\underline{v}^{\mathrm{T}} \underline{x} - b_v = -b_v > 0$ - both constraints are in fact violated at this solution. Hence, $\underline{x} = 0$ is not a feasible minima of the constrained problem.

$A(\underline{x}^*) = \{u\}$: In this case, the necessary condition becomes

$$\nabla f(\underline{x}) + \mu_u \nabla g_u(\underline{x}) = 0 \tag{109}$$

for $\mu_u \geqslant 0$, or or

$$2\underline{x} + \mu_u \underline{u} = 0. \tag{110}$$

This has a unique solution (for a given μ_u) of $\underline{x} = -(\mu_u)/2\underline{u}$. As u is the active constraint, we must have $g_u(\underline{x}) = 0$, which can be utilized to solve for μ_u and give a final solution of

$$\underline{x}_u^* = \frac{b_u}{||\underline{u}||^2}\underline{u}. \tag{111}$$

This will be a feasible minimum, as long as v is inactive, i.e., $g_v(\underline{x}_u^*) < 0$ or

$$\frac{\underline{v}^{\mathrm{T}}\underline{u}}{||\underline{u}||^2}b_u < b_v. \tag{112}$$

If it is feasible, it will yield a function value of $f(\underline{x}_u^*) = b_u$.

$A(\underline{x}^*) = \{v\}$: By the symmetry of the problem, we may apply the analysis of the previous case to yield a solution

$$\underline{x}_v^* = \frac{b_v}{||\underline{v}||^2}\underline{v}, \tag{113}$$

which will be feasible if

$$\frac{\underline{u}^{\mathrm{T}}\underline{v}}{||\underline{v}||^2}b_v < b_u. \tag{114}$$

If it is feasible, it will yield a function value of $f(\underline{x}_v^*) = b_v$.

$A(\underline{x}^*) = \{u, v\}$: In this case, we take both u and v constraints to be active, i.e., $g_v(\underline{x}) = 0$ and $g_u(\underline{x}) = 0$. Note we have that $\nabla g_u(\underline{x}) = \underline{u}$ and $\nabla g_v(\underline{x}) = \underline{v}$ are linearly independent, and hence the regularity condition is satisfied.

In this case we have

$$\nabla f(\underline{x}) + \mu_u \nabla g_u(\underline{x}) + \mu_v \nabla g_v(\underline{x}) = 0 \tag{115}$$

for $\mu_u, \mu_v \geqslant 0$, or

$$2\underline{x} + \mu_u\underline{u} + \mu_v\underline{v} = 0. \tag{116}$$

The above, along with the constraints that $g_u(\underline{x}) = 0, g_v(\underline{x}) = 0$, may be used to solve the entire system to yield

$$\begin{aligned} \mu_u &= \frac{-2(b_v\underline{u}^{\mathrm{T}}\underline{v} - b_u||\underline{v}||^2)}{(\underline{u}^{\mathrm{T}}\underline{v})^2 - ||\underline{u}||^2||\underline{v}||^2} \\ \mu_v &= \frac{-2(-b_v||\underline{u}||^2 + b_u\underline{u}^{\mathrm{T}}\underline{v})}{(\underline{u}^{\mathrm{T}}\underline{v})^2 - ||\underline{u}||^2||\underline{v}||^2} \\ \underline{x}_{uv}^* &= \frac{(b_v\underline{u}^{\mathrm{T}}\underline{v} - b_u||\underline{v}||^2)}{(\underline{u}^{\mathrm{T}}\underline{v})^2 - ||\underline{u}||^2||\underline{v}||^2}\underline{u} + \frac{(-b_v||\underline{u}||^2 + b_u\underline{u}^{\mathrm{T}}\underline{v})}{(\underline{u}^{\mathrm{T}}\underline{v})^2 - ||\underline{u}||^2||\underline{v}||^2}\underline{v}, \end{aligned} \tag{117}$$

which can be shown to be feasible (under the constraints $\mu_u, \mu_v \geqslant 0$) so long as

$$b_v\frac{\underline{u}^{\mathrm{T}}\underline{v}}{||\underline{v}||^2} \geqslant b_u \text{ and } b_u\frac{\underline{u}^{\mathrm{T}}\underline{v}}{||\underline{u}||^2} \geqslant b_v. \tag{118}$$

The above solution can be shown to yield a function value of

$$f(\underline{x}_{uv}^*) = \frac{||b_v\underline{u} - b_u\underline{v}||^2}{||\underline{u}||^2||\underline{v}||^2 - (\underline{u}^{\mathrm{T}}\underline{v})^2}. \tag{119}$$

Which of these is the true minimum (and it can be shown via the second order conditions, or geometric intuition about the objective function) will depend on the specific parameters involved.

6.4 Example Problems

1 In the solution of Eq. (84), it is concluded via Lagrange multipliers that the minimizer must be some scalar multiple of $\underline{u}$, i.e., the minimizer must be a vector in the same or opposite direction to $\underline{u}$. Argue that this must be true, from the geometric interpretation of the problem. Additionally; it was stated that the solution in Eq. (89) must be the minimizer. Prove that this is true, and that this point is not actually a maximizer.

2 Let Q be a real symmetric matrix with $\lambda_{\min}, \lambda_{\max}$ as the smallest and largest eigenvalues, respectively. Using Lagrange multipliers, prove the following identity:

$$\lambda_{\min}||\underline{x}||^2 \leqslant \underline{x}^{\mathrm{T}} Q \underline{x} \leqslant \lambda_{\max}||\underline{x}||^2. \tag{120}$$

Hint: What is the largest and smallest value of $\underline{x}^T Q \underline{x}$ for a given Q? Note that Q as an orthonormal basis of eigenvectors. Also, it suffices to consider the case of $||\underline{x}|| = 1$ - why?

3 Let $\{\alpha_i\}$ be a set of constants such that $\alpha_i > 0$ and $\sum_i \alpha_i = 1$. Use Lagrange multipliers to solve

$$\begin{aligned} \text{minimize } & \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n \\ \text{s.t. } & \prod_{i=1}^{n} x_i^{\alpha_i} = 1 \\ & x_i \geqslant 0 \text{ for all } i\ . \end{aligned} \tag{121}$$

Hint: Consider the change of variables $x_i = e^{y_i}$ for $y_i \in \mathbb{R}$. What can you conclude about the case of $x_i = 0$ for some i?

4 Solve the following problem:

$$\begin{aligned} \text{minimize } & x_1 - 2x_2 \\ \text{s.t. } & x_1^2 + x_2^2 \leqslant 4 \\ & (x_1 - 3)^2 + x_2^2 \leqslant 4. \end{aligned} \tag{122}$$

What is the geometry of the constraint set?

7 Analytic Based Algorithms

In this section, we consider algorithms based on the idea of analytically defined constraints. We consider two particular cases, equality constraints:

$$X = \{\underline{x} \in \mathbb{R}^n : \underline{h}(\underline{x}) = 0\}, \tag{123}$$

and inequality constraints

$$X = \{\underline{x} \in \mathbb{R}^n : \underline{g}(\underline{x}) = 0\}. \tag{124}$$

The **Barrier** or **Interior Point Methods** are better suited to inequality constraint problems, while **Penalty Methods** are better suited to equality constraint problems. The concepts behind each method may potentially be combined and applied to mixed constraint problems, as appropriate.

In general, the theme of this section is solving constrained optimization problems by reducing them to a sequence of unconstrained optimization problems, whose solutions $\{\underline{x}_k^*\}$ (hopefully) converge to the solution of the constrained problem, $\underline{x}^*$.

7.1 Barrier/Interior Point Methods

Let $X = \{\underline{x} \in \mathbb{R}^n : g_j(\underline{x}) \leqslant 0 \text{ for } j = 1, \ldots r\}$, or more compactly, $\underline{g}(\underline{x}) \leqslant 0$. Let S be defined as the *interior* of X, that is, $S \subset X$ such that

$$S = \{\underline{x} \in \mathbb{R}^n : g_j(\underline{x}) < 0 \text{ for } j = 1, \ldots, r\}. \tag{125}$$

The central assumption (so as to 'jumpstart' the method) is that the interior is non-empty, so that there is some $\underline{x}_0 \in S$. **Barrier Methods** proceed by augmenting the objective function f to mathematically create a barrier at the boundary of X, so that unconstrained optimization methods are unable to step outside the constraint set. Consider the objective function

$$f(\underline{x}) + \varepsilon B(\underline{x}), \tag{126}$$

where $\varepsilon > 0$ is very small, and $B(\underline{x})$ increases to $+\infty$ at the boundary of X. Over the interior of X, choosing ε and B appropriately, this augmented objective function will be approximately equal to f, and hence a minimizer of the augmented objective function will (hopefully) approximately minimize f. The important concept however is the following; starting from a point in the interior of X (i.e., $\underline{x}_0$), an unconstrained minimization algorithm applied to the augmented objective function will never step outside the constraint set X - stepping outside of it will always lead to an increase in the objective function, to infinity. Hence, if we begin in the interior of X, we may apply unconstrained minimization methods and remain in the interior of X, satisfying the constraints.

The two most common barrier functions are the *harmonic* and the *logarithmic* barrier functions:

- **Harmonic Barrier:**
$$B(\underline{x}) = \sum_{j=1}^{r} \frac{1}{-g_j(\underline{x})}. \tag{127}$$

- **Logarithmic Barrier:**
$$B(\underline{x}) = -\sum_{j=1}^{r} \ln(-g_j(\underline{x})). \tag{128}$$

The logarithmic barrier function has proven to be particularly useful in application to Linear Programming problems; the first polynomial-time LP algorithms were based on related techniques.

This leads to the following construction: Let $\{\varepsilon_k\}$ be a decreasing sequence of positive values such that $\varepsilon_k \to 0$ as $k \to \infty$. For a given barrier function B, we may construct a sequence of augmented objective functions

$$f_k(\underline{x}) = f(\underline{x}) + \varepsilon_k B(\underline{x}). \tag{129}$$

We then define the sequence $\{\underline{x}_k^*\}$ (with $\underline{x}_0^* = \underline{x}_0$) to be the sequence of *unconstrained global minima* of f_k, i.e., for $k > 0$,

$$\underline{x}_k^* = \arg\min_{\underline{x}\in\mathbb{R}^n} f_k(\underline{x}). \tag{130}$$

Note, from a computational or algorithmic perspective, when solving for $\underline{x}_{k+1}^*$, it can frequently be useful to use $\underline{x}_k^*$ as the initial guess for minimizing f_{k+1}. This is what is known as a ***warm start****.*

Ideally, as $\varepsilon_k \to 0$, the barrier function will have less and less of an impact, and hence the minima of f_k should increasingly correspond to the minima of f over the constraint set X, and hence (hopefully), $\underline{x}_k^*$ should converge to some minimum $\underline{x}^*$ of f over X.

We do have the following result:

Theorem 10 *If the sequence $\{\underline{x}_k^*\}$ is defined to be the successive, unconstrained global minima of f_k, then if $\underline{x}_k^* \to \underline{x}^*$ as $k \to \infty$, we have that $\underline{x}^*$ is a global minimum of f over the constraint set X.*

Of course, computing $\underline{x}_k^*$ as a global minimum of f_k is ambitious, based on established unconstrained minimization algorithms; in general, we may only hope to compute, approximately, local minima. However, if we have additional results on the structure of f, for instance f as strictly convex over X, we may frequently conclude that the minima of f_k are unique and global.

7.2 Penalty Methods

In the case of barrier methods, the constraint set was (analytically) defined to be a volume in $\mathbb{R}^n$, and the barrier method proceeded by moving through this volume, minimizing a sequence of objective functions that approximated the primary objective function. In the case of equality constraints, however,

$$X = \{\underline{x} \in \mathbb{R}^n : \underline{h}(\underline{x}) = 0\}, \tag{131}$$

it can be difficult to restrict 'motion' in such a way as to remain inside the constraint set (indeed, finding an initial point in the constraint set may be difficult in itself).

Therefore, if Barrier Methods can be thought of in terms of working from the inside of the constraint set out, Penalty Methods are constructed to work from the outside in, producing a sequence of points outside the constraint set that converges to a minimum of f inside the constraint set. In particular, we augment the objective function with a *penalty* or *cost* for violating the constraints.

Let $P(\delta)$ be a positive, monotonically increasing function for $\delta \geqslant 0$. In this case, for $c > 0$ we may consider the augmented objective function over all $\mathbb{R}^n$

$$f(\underline{x}) + cP(||\underline{h}(\underline{x})||). \tag{132}$$

The farther a given $\underline{x}$ is from satisfying the constraints, i.e., the larger $||\underline{h}(\underline{x})||$, the larger the value of the augmented objective function. This penalty function creates a sort of pressure towards satisfying the constraint, when unconstrained minimization algorithms are applied. The larger the value of c, the more unconstrained minimization algorithms will be driven toward X.

The most common penalty function is the quadratic penalty, $P(\delta) = \frac{1}{2}\delta^2$. In some contexts however, depending on how f scales with h and how close you start to a local minimum, it may be useful to consider alternative penalty functions as well, such as $P(\delta) = e^{\delta}$.

We therefore introduce the following sequence of unconstrained minimization problems: Let $\{c_k\}$ be a positive, increasing sequence, and define the augmented objective function

$$f_k(\underline{x}) = f(\underline{x}) + c_k P(||\underline{h}(\underline{x})||). \tag{133}$$

We then define the sequence $\{\underline{x}_k^*\}$ to be the sequence of global, unconstrained minimizers of the augmented objective functions:

$$\underline{x}_k^* = \arg\min_{\underline{x}\in\mathbb{R}^n} f_k(\underline{x}). \tag{134}$$

Note again, it may be useful to recycle $\underline{x}_k^*$ as an initial guess for the minimum of f_{k+1}, as a warm start.

Ideally, as c_k increases, a larger and larger cost is paid for not satisfying the constraints; this (hopefully) creates a pressure drawing the $\underline{x}_k^*$ toward the constraint set X, such that $\{\underline{x}_k^*\}$ converges to a minimum of f over X. Indeed, we have the following theorem:

Theorem 11 *Taking $P(\delta) = (1/2)\delta^2$ and $c_k \to \infty$ as $k \to \infty$: If the sequence $\{\underline{x}_k^*\}$ is defined to be the successive, unconstrained global minima of f_k, then if $\underline{x}_k^* \to \underline{x}^*$ as $k \to \infty$, we have that $\underline{x}^*$ is a global minimum of f over the constraint set X.*

This should give some confidence in the utility of the method. However, again the result is predicated on the ability to identify unconstrained global minima of each successive augmented objective function. In reality, we must be content with approximating local minima (given finite computational resources). But stronger structural assumptions on f, such as convexity, may be useful here.

7.3 The Augmented Lagrangian and Multiplier Methods

As an extension of general penalty methods, we have the **augmented Lagrangian method**. Consider the following function:

$$L_c(\underline{x}, \underline{\lambda}) = f(\underline{x}) + \underline{\lambda}^{\mathrm{T}}\underline{h}(\underline{x}) + \frac{1}{2}c||\underline{h}(\underline{x})||^2. \tag{135}$$

For a given $\underline{\lambda}, c$, consider the problem of minimizing (in an unconstrained fashion) $L_c(\underline{x}, \underline{\lambda})$. If c is large relative to the scale of $\underline{\lambda}$, the penalty term will dominate the additional Lagrangian term, and hence minimization of $L_c(\underline{x}, \underline{\lambda})$ will approximate minimization of f subject to the constraint $\underline{h}(\underline{x}) = 0$.

Indeed, we have the following theorem:

Theorem 12 *Let c_k be a positive, increasing sequence with $c_k \to \infty$ as $k \to \infty$. Let $\{\underline{\lambda}_k\}$ be a bounded sequence. Let $\{\underline{x}_k^*\}$ be the sequence of global minimizers:*

$$\underline{x}_k^* = \arg\min_{\underline{x}\in\mathbb{R}^n} L_{c_k}(\underline{x}, \underline{\lambda}_k). \tag{136}$$

In this case, if $\underline{x}_k^* \to \underline{x}^*$, then $\underline{x}^*$ represents a global minimum of f over X.

The augmented Lagrangian method presents two main advantages to the pure penalty method: i) a theoretical basis for approximate minimization of the augmented objective functions; ii) convergence can be dramatically improved (see associated Mathematica workbook) by choosing the sequence $\underline{\lambda}_k$ to successively approximate a Lagrange multiplier $\underline{\lambda}^*$. We have the following results:

Theorem 13 *Let $\underline{x}_k$ satisfy*

$$||\nabla_x L_{c_k}(\underline{x}_k, \underline{\lambda}_k)|| \leqslant \varepsilon_k, \tag{137}$$

where $\{\lambda_k\}$ is bounded, $\{\underline{c}_k\}$ is positive, increasing to infinity, and $\{\varepsilon_k\}$ is positive, decreasing to 0. *Note, $\underline{x}_k$ may be viewed as an approximate minimum of the augmented Lagrangian, as*

$$\nabla_x L_{c_k}(\underline{x}_k^*, \underline{\lambda}_k) = 0. \tag{138}$$

If $\{\underline{x}_k\}$ converges to $\underline{x}^*$ such that $\nabla \underline{h}(\underline{x}^*)$ has rank m (i.e., the point $\underline{x}^*$ is regular), then

$$\underline{\lambda}_k + c_k \underline{h}(\underline{x}_k) \to \underline{\lambda}^* \tag{139}$$

where

$$\nabla f(\underline{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla h_i(\underline{x}^*) = 0 \textit{ and } \underline{h}(\underline{x}^*) = 0. \tag{140}$$

Note, the above result suggests that under the appropriate conditions, it is not necessary to compute the global minimum of each augmented Lagrangian - it suffices to consider sufficient approximations to local minima. Any limit point that results from this will satisfy the first order Lagrange necessary conditions for a minimum. Obviously the choice of ε_k will affect the rate of convergence, as will the c_k and the choice of $\underline{\lambda}_k$.

Additionally, this suggests a scheme for successively estimating the Lagrange multiplier corresponding to a constrained minimum, the so called Method of Multipliers: take

$$\underline{\lambda}_{k+1} = \underline{\lambda}_k + c_k \underline{h}(\underline{x}_k) \tag{141}$$

as long as the resulting vector is not too large in magnitude (consider taking $\underline{\lambda}_{k+1} = \underline{\lambda}_k$ in such a case). It can be shown in many cases that if $\underline{\lambda}^*$ may be effectively approximated by such a scheme, convergence can be guaranteed without taking c_k to ∞; this can stabilize many numerical algorithms. In general, as c_k increases, the rate of convergence of $\underline{\lambda}_k$ increases as well.

In general, if the $\underline{x}_k$ are sufficiently accurate estimates of the local minima of the augmented Lagrangian, it can be shown that the rate of convergence in the errors, $||\underline{x}_k - \underline{x}^*||$ and $||\underline{\lambda}_k - \underline{\lambda}^*||$, is linear, improving to superlinear if c_k increases to infinity.

Some important computational considerations: The cost factor c_k should become sufficiently large eventually - however starting it too large can bias the algorithm towards minimizing the penalty function alone, rather than attempting to minimize the objective function f simultaneously, which will lead to poor convergence properties. Similarly, if c_k increases too slowly, this can lead to poor convergence of the multipliers $\underline{\lambda}_k$. Experimentation may be useful for judging appropriate parameter values. If the unconstrained minimization of the augmented Lagrangian is performed with a powerful enough method such as Newton's method, a frequent scheme is to increase the value c_k by a factor of $\beta \in [5, 10]$ in each step; less effective means may require a slower increase in c_k. Another technique is to tie the increase in c_k to the reduction in the constraint violated $||\underline{h}(\underline{x}_k)||/||\underline{h}(\underline{x}_{k-1})||$, only increasing c_k if a sufficient reduction in constraint violation has *not* been achieved.

7.4 Example Problems

1 What are the gradients of the logarithmic and harmonic barrier functions? If the g_j are convex, are these barrier functions convex?

8 Duality

In the previous sections, we examined how the motivating n-dimensional constrained minimization problem (for instance, in the equality case)

$$\begin{aligned} \text{minimize } & f(\underline{x}) \\ \text{s.t. } & \underline{h}(\underline{x}) = 0, \\ & \underline{x} \in \mathbb{R}^n \end{aligned} \tag{142}$$

could be approached by embedding it in a larger $(n+m)$-dimensional problem, centered on the Lagrangian function

$$L(\underline{x}, \underline{\lambda}) = f(\underline{x}) + \underline{\lambda}^{\mathrm{T}} \underline{h}(\underline{x}), \tag{143}$$

and finding the solutions $(\underline{x}^*, \underline{\lambda}^*)$ to the system of Lagrange multiplier equations.

Taking this one step further, the problem can be extended by a sort of symmetry into an optimization *only* over the $\underline{\lambda}$-variable (the ‘dual’ variables), as an m-dimensional maximization problem. This is the essence of *duality*, which is the subject of this section. This has many important implications and applications, both in terms of understanding the underlying structure of optimization problems, and potentially granting many computational advantages as well, for instance in simplifying the problem or admitting new algorithmic approaches. In many cases, for instance when $r < n$, the ‘dual’ optimization problem may be much more computationally tractable, and many problems of practical concern are structurally much simpler in the dual formulation. Additionally, the treatment presented here will allow extension and application of the theory to non-differentiable or even discrete cases.

Consider, for the moment, the more generalized problem, subject for now only to equality constraints:

$$\begin{aligned} \text{minimize } & f(\underline{x}) \\ \text{s.t. } & \underline{h}(\underline{x}) = 0, \\ & \underline{x} \in X, \end{aligned} \tag{144}$$

where X is some domain of interest (potentially a subset of $\mathbb{R}^n$, potentially $\mathbb{R}^n$, or potentially discrete).

Denote the *primal optimum value* as

$$f^* = \inf_{\underline{x} \in X, \underline{h}(\underline{x})=0} f(\underline{x}). \tag{145}$$

We utilize the infimum in the above for complete generality, to account for the case that the minimum is not actually realized. In the remainder of the section we will assume that *there are feasible solution* $\underline{x}$, and that the optimum value is *bounded*, i.e., $-\infty < f^* < \infty$.

The definition of Lagrange multipliers depended explicitly on a notion of differentiability of the objective and constraint functions. Here, we generalize this to the concept of *geometric multipliers*: a geometric multiplier $\underline{\lambda}^*$ if

$$f^* = \inf_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}^*). \tag{146}$$

Note, this is a natural extension of the notion of Lagrange multipliers, as they implicitly satisfy $f(\underline{x}^*) = L(\underline{x}^*, \underline{\lambda}^*)$ due to the nature of the equality constraints. But again, we make no assumption concerning differentiability or even the underlying domain here - in the case of differentiability, it can be shown that Lagrange multipliers and geometric multipliers correspond for optimal $\underline{x}^*$.

Taking a geometric multiplier $\underline{\lambda}^*$ as known, however, all optimal solutions to the primal problem can be recovered via *unconstrained* optimization, minimizing $L(\underline{x}, \underline{\lambda}^*)$ over all $\underline{x}$. Namely, we have the following result:

Proposition 2 *Let $\underline{\lambda}^*$ be a geometric multiplier. A point $\underline{x}^*$ is a global minimum of the primal problem **iff** $\underline{x}^*$ is feasible and $\underline{x}^* \in \arg\min_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}^*)$.*

Proof. Let $\underline{x}^*$ be a global minimum of the primal problem, and hence feasible with respect to the constraints, i.e., $\underline{h}(\underline{x}^*) = 0$. Then

$$f^* = f(\underline{x}^*) = f(\underline{x}^*) + \underline{\lambda}^* \underline{h}(\underline{x}^*) = L(\underline{x}^*, \underline{\lambda}^*) \geqslant \inf_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}^*) = f^*, \tag{147}$$

the last equality following by the definition of a geometric multiplier. It follows then that equality is realized at every step, and $L(\underline{x}^*, \underline{\lambda}^*) = \min_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}^*)$.

To complete the proof in the other direction, note that taking $\underline{x}^*$ as feasible, and $\underline{x}^* \in \arg\min_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}^*)$, we have that $f(\underline{x}^*) = L(\underline{x}^*, \underline{\lambda}^*) = \min_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}^*) = f^*$. Hence, $\underline{x}^*$ is a global minimum of the primal problem. □

Note importantly: no assumptions regarding differentiability or the underlying domain were made.

This indicates the importance of geometric multipliers $\underline{\lambda}^*$ as objects worth finding in themselves. This motivates the definition of the **dual function**:

$$q(\underline{\lambda}) = \inf_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}), \tag{148}$$

and the **dual problem**:

$$\begin{aligned} \text{maximize } & q(\underline{\lambda}) \\ \text{s.t. } & \underline{\lambda} \in \mathbb{R}^m. \end{aligned} \tag{149}$$

It will be shown that solutions $\underline{\lambda}^*$ (optimal *dual* solutions) essentially yield geometric multipliers. This will be made precise presently. Note, the dual problem is taken to be *unconstrained*. It can be constrained however, in the following way: noting that q may sometimes be infinite, define the domain of the dual function as

$$D = \{\underline{\lambda} \in \mathbb{R}^m : q(\underline{\lambda}) > -\infty\}. \tag{150}$$

In this case, the dual problem still has important analytic properties, namely:

Proposition 3 *The dual function q is concave over its domain D, which is a convex subset of $\mathbb{R}^n$.*

The relationship between the dual problem and the primal problem (and hence the utility of the dual) is summarized in the following relation: let $\underline{x}' \in X$ be feasible, i.e., $\underline{h}(\underline{x}') = 0$. Then,

$$q(\underline{\lambda}) = \inf_{\underline{x} \in X} L(\underline{x}, \underline{\lambda}) \leqslant f(\underline{x}') + \underline{\lambda}^{\mathrm{T}} \underline{h}(\underline{x}') = f(\underline{x}') \leqslant f^*. \tag{151}$$

Taking the supremum of the above over $\underline{\lambda}$ (denote q^* on the left), we have the following result:

Theorem 14 (The Weak Duality Theorem and Characterization of Geometric Multipliers) *We have the following, always*

$$q^* \leqslant f^*. \tag{152}$$

Additionally, if there is no ***duality gap****, i.e., $q^* = f^*$, then any geometric multiplier is a dual optimal solution and any optimal dual solution is a geometric multiplier. If there is a duality gap, there are no geometric multipliers.*

Proof. The first result stems, as stated, from taking the supremum over $\underline{\lambda}^*$ above. To prove the second part, note that in the case that there is no duality gap, we have for any geometric multiplier $f^* = q(\underline{\lambda}^*) \leqslant q^*$ which, by the initial result above, implies $q(\underline{\lambda}^*) = q^*$ and hence it is an optimal dual solution. Similarly, if there is a duality gap, i.e., $q^* < f^*$, no such $\underline{\lambda}^*$ satisfying $f^* = q(\underline{\lambda}^*)$ can exist.

The weak duality theorem above is useful in the sense that the dual problem always provides a lower bound on the value of the primal problem. When there is no duality gap, however, the primal problem can be exchanged for the dual problem, as they yield equivalent optimal values. This is particularly useful when the dual problem is of particularly simple structure.

We pause here to observe that the above results can be extended naturally and almost immediately to the case of *inequality* constraints $\underline{g}(\underline{x}) \leqslant 0$, extending the Lagrangian to

$$L(\underline{x}, \underline{\lambda}, \underline{\mu}) = f(\underline{x}) + \underline{\lambda}^{\mathrm{T}} \underline{h}(\underline{x}) + \underline{\mu}^{\mathrm{T}} \underline{g}(\underline{x}), \tag{153}$$

and extending the definition of geometric multipliers appropriately, now with the added constraint that $\underline{\mu} \geqslant 0$, and the introduction of complementary slackness conditions (as seen in the corresponding results for Lagrange multipliers) that $\mu_j^* g_j(\underline{x}^*) = 0$ for $j = 1, \ldots r$. The dual function is naturally extended, as is the dual problem, with the additional constraint on $\underline{\mu}$. The weak duality theorem remains as is.

The following two cases exhibit no duality gap, and hence the primal problem may be exchanged for the dual problem:

- Let f be convex over $\mathbb{R}^n$, and let the constraints be linear, i.e., X as polyhedral in $\mathbb{R}^n$. Additionally, take f^* as finite. In this case, there is no duality gap, and there exists at least one geometric multiplier.
- Let f be convex quadratic and let the constraints be linear, i.e., X as polyhedral in $\mathbb{R}^n$. Additionally, take f^* as finite. In this case, thre is no duality gap, and primal optimal and dual optimal solutions exist.

We consider two examples of the above cases, in which we may utilize the dual problem to solve or simplify the primal problem.

Example One: Consider the problem of projection onto a linear subspace, i.e., for fixed $\underline{z} \in \mathbb{R}^n$, matrix $A \in \mathbb{R}^{m\times n}$:

$$\begin{aligned} &\min \,||\underline{z}-\underline{x}||^2 \\ &\text{s.t. } A\underline{x} = 0. \end{aligned} \tag{154}$$

In this case, the Lagrangian is given by $L(\underline{x},\underline{\lambda}) = ||\underline{z}-\underline{x}||^2 + \underline{\lambda}^{\mathrm{T}}A\underline{x}$. Hence we may express the dual function (as an unconstrained optimization problem):

$$\begin{aligned} q(\underline{\lambda}) &= \inf_{\underline{x}\in\mathbb{R}^n}\left[||\underline{z}-\underline{x}||^2 + \underline{\lambda}^{\mathrm{T}}A\underline{x}\right] \\ &= \inf_{\underline{x}\in\mathbb{R}^n}\left[\underline{x}^{\mathrm{T}}\underline{x} + \left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right)^{\mathrm{T}}\underline{x}\right] + ||\underline{z}||^2 \end{aligned} \tag{155}$$

Note, in the usual way, the above is realized when $2\underline{x} + (A^{\mathrm{T}}\underline{\lambda}-2\underline{z}) = 0$, or

$$\begin{aligned} \underline{x}^* &= -\frac{1}{2}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right) \\ q(\underline{\lambda}) &= \frac{1}{4}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right)^{\mathrm{T}}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right) - \frac{1}{2}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right)^{\mathrm{T}}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right) + ||\underline{z}||^2 \\ &= -\frac{1}{4}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right)^{\mathrm{T}}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right) + ||\underline{z}||^2 \\ &= -\frac{1}{4}\underline{\lambda}^{\mathrm{T}}AA^{\mathrm{T}}\underline{\lambda} + \underline{z}^{\mathrm{T}}A^{\mathrm{T}}\underline{\lambda} \end{aligned} \tag{156}$$

Hence, the *dual* problem may equivalently be expressed as the unconstrained problem

$$\begin{aligned} &\max \; -\frac{1}{4}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right)^{\mathrm{T}}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right) + ||\underline{z}||^2 \\ &\text{s.t. } \underline{\lambda} \in \mathbb{R}^m. \end{aligned} \tag{157}$$

In terms of solving for the solution $\underline{\lambda}^*$, note that the $+||\underline{z}||^2$ is constant, and thus may be ignored, and we may dispense with the $-1/4$ factor to turn it into again a minimization problem,

$$\begin{aligned} &\min \; \left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right)^{\mathrm{T}}\left(A^{\mathrm{T}}\underline{\lambda}-2\underline{z}\right) \\ &\text{s.t. } \underline{\lambda} \in \mathbb{R}^m. \end{aligned} \tag{158}$$

As this is an unconstrained problem, we may apply the usual techniques to it, in particular observing that (expanding and taking the gradient, setting it equal to 0), any solution $\underline{\lambda}^*$ must satisfy

$$AA^{\mathrm{T}}\underline{\lambda}^* = 2A\underline{z}. \tag{159}$$

Example Two: Consider the following problem, for symmetric, positive definite Q, and fixed $\underline{v}$:

$$\begin{aligned} \min\ & \frac{1}{2}\underline{x}^{\mathrm{T}}Q\underline{x} \\ \text{s.t. } & \underline{v}^{\mathrm{T}}\underline{x} \geqslant 1. \end{aligned} \tag{160}$$

The Lagrangian of the above is given by $L(\underline{x},\mu) = (1/2)\underline{x}^{\mathrm{T}}Q\underline{x} + \mu(1-\underline{v}^{\mathrm{T}}\underline{x})$. This leads to a dual function of

$$q(\mu) = \inf_{\underline{x}\in\mathbb{R}^n}\left[(1/2)\underline{x}^{\mathrm{T}}Q\underline{x} + \mu\left(1-\underline{v}^{\mathrm{T}}\underline{x}\right)\right], \tag{161}$$

which is maximized taking $Q\underline{x}^* - \mu\underline{v} = 0$, or $\underline{x}^* = \mu Q^{-1}\underline{v}$:

$$\begin{aligned} q(\mu) &= \left[(1/2)\mu^2\underline{v}^{\mathrm{T}}Q^{-1}\underline{v} + \mu\left(1-\mu\underline{v}^{\mathrm{T}}Q^{-1}\underline{v}\right)\right] \\ &= \mu - \frac{1}{2}\mu^2\underline{v}^{\mathrm{T}}Q^{-1}\underline{v}. \end{aligned} \tag{162}$$

The dual problem is then immediately

$$\begin{aligned} \max\ & \mu - \frac{1}{2}\mu^2\underline{v}^{\mathrm{T}}Q^{-1}\underline{v} \\ \text{s.t. } & \mu \geqslant 0. \end{aligned} \tag{163}$$

Note importantly: because of the structure of the constraints, by exchanging from the primal to the dual we have traded an n-dimensional constrained optimization problem for a 1-dimensional constrained optimization problem. This is a dramatic simplification.

Note that the unconstrained maximum of the above objective function occurs when $1 = \mu^*\underline{v}^{\mathrm{T}}Q^{-1}\underline{v}$ or

$$\mu^* = \frac{1}{\underline{v}^{\mathrm{T}}Q^{-1}\underline{v}} > 0. \tag{164}$$

As the above solution is strictly positive, it is also the constrained maximum. Substituting, this yields an optimal solution of

$$\underline{x}^* = \frac{Q^{-1}\underline{v}}{\underline{v}^{\mathrm{T}}Q^{-1}\underline{v}} \tag{165}$$

and an optimal value of

$$\frac{1}{2}\frac{1}{\left(\underline{v}^{\mathrm{T}}Q^{-1}\underline{v}\right)}. \tag{166}$$

The following results, known as the *Slater Constraint Qualification* give a condition in a general case for there being no duality gap, and the existence of a geometric multiplier. In particular, in the inequality constraint case:

$$\begin{aligned} \min\ & f(\underline{x}) \\ \text{s.t. } & \underline{g}(\underline{x}) \leqslant 0 \\ & \underline{x} \in X, \end{aligned} \tag{167}$$

with the following assumptions, that there exist feasible points $\underline{x} \in X$ with $\underline{g}(\underline{x}) \leqslant 0$, the optimal value f^* is finite, X is a convex subset of $\mathbb{R}^n$, and the functions f and $\{g_j\}$ are convex over X. In addition, (and this is the Slater qualification), there exists an *interior* point $\underline{x}' \in X$ such that $\underline{g}(\underline{x}') < 0$.

In this case, we have that there is no duality gap, and there exists at least one geometric multiplier.

8.1 Separable Problems

It is often the case, particularly in large data applications, that we can consider the contribution to both the objective function and the various constraints separately for instance for different subsets of data. For $i = 1, \ldots, N$, let $\underline{x}_i$ represent the variable associated with 'block' i, and f_i, $\underline{h}_i$, $\underline{g}_i$ the objective function and constraint contributions of this block, and consider the problem

$$\begin{aligned} \min \; & \sum_{i=1}^{N} f_i(\underline{x}_i) \\ \text{s.t.} \; & \sum_{i=1}^{N} \underline{h}_i(\underline{x}_i) = 0 \\ & \sum_{i=1}^{N} \underline{g}_i(\underline{x}_i) \leqslant 0. \end{aligned} \tag{168}$$

In this case, we may express the Lagrangian as

$$L(\underline{x}_1, \ldots, \underline{x}_N, \underline{\lambda}_1, \ldots, \underline{\lambda}_N, \underline{\mu}_1, \ldots, \underline{\mu}_N) = \sum_{i=1}^{N} \Big(f_i(\underline{x}_i) + \underline{\lambda}_i^{\mathrm{T}} \underline{h}_i(\underline{x}_i) + \underline{\mu}^{\mathrm{T}} \underline{g}_i(\underline{x}_i) \Big). \tag{169}$$

The dual function then becomes

$$\begin{aligned} q\Big(\{\underline{\lambda}_i\}, \{\underline{\mu}_i\}\Big) &= \inf_{\{\underline{x}_i \in X_i\}} \left[\sum_{i=1}^{N} \Big(f_i(\underline{x}_i) + \underline{\lambda}_i^{\mathrm{T}} \underline{h}_i(\underline{x}_i) + \underline{\mu}^{\mathrm{T}} \underline{g}_i(\underline{x}_i) \Big) \right] \\ &= \sum_{i=1}^{N} \inf_{\underline{x}_i \in X_i} \Big(f_i(\underline{x}_i) + \underline{\lambda}_i^{\mathrm{T}} \underline{h}_i(\underline{x}_i) + \underline{\mu}^{\mathrm{T}} \underline{g}_i(\underline{x}_i) \Big) \\ &= \sum_{i=1}^{N} q_i(\underline{\lambda}_i, \underline{\mu}_i), \end{aligned} \tag{170}$$

where

$$q_i(\underline{\lambda}_i, \underline{\mu}_i) = \inf_{\underline{x}_i \in X_i} \left[f_i(\underline{x}_i) + \underline{\lambda}_i^{\mathrm{T}} \underline{h}_i(\underline{x}_i) + \underline{\mu}^{\mathrm{T}} \underline{g}_i(\underline{x}_i) \right]. \tag{171}$$

In this way, the problem of maximizing $q\Big(\{\underline{\lambda}_i\}, \{\underline{\mu}_i\}\Big)$ subject to the constraints $\{\underline{\mu}_i \geqslant 0\}$ can be exchanged for N *separate* problems of the form

$$\begin{aligned} \max \; & q_i(\underline{\lambda}_i, \underline{\mu}_i) \\ \text{s.t.} \; & \underline{\mu}_i \geqslant 0 \\ & \underline{\lambda}_i \in \mathbb{R}^{m_i}. \end{aligned} \tag{172}$$

This can represent a significant dimension reduction and hence significant computational savings, for problems of this form.

8.2 Example Problems

1 Prove Prop. 3. *Hint: Consider* $L(\underline{x}, \alpha\underline{\lambda} + (1-\alpha)\underline{\lambda}')$.

2 Formulate the primal and dual problems for linear programming within the context of the material in this section.

3 Consider the problem of minimizing e^y subject to $\sqrt{x^2+y^2} \leqslant x$, for $(x,y) \in \mathbb{R}^2$. Calculate f^* and q^*. Is there a duality gap?

A The Extreme Value Theorem

Taking $X \subset \mathbb{R}^n$ to be *closed and bounded*, we have via the Heine-Borel Theorem that X is *compact*. What this means practically is that given any sequence $\underline{x}_1, \underline{x}_2, \ldots \in X$ there is a subsequence $\underline{x}_{k_1}, \underline{x}_{k_2}, \ldots$ such that $\underline{x}_{k_n}$ converges to some $\underline{x}' \in X$.

Let $f^* = \inf_{\underline{x}\in X} f(\underline{x})$, i.e., $f^* \leqslant f(\underline{x})$ for all $\underline{x} \in X$. The goal is to identify some $\underline{x}^* \in X$ such that $f^* = f(\underline{x}^*)$. To do so, consider defining a sequence $\{\underline{x}_k\}$ in such a way that for any $k \geqslant 1$,

$$f(\underline{x}_k) - f^* \leqslant \frac{1}{k}. \tag{173}$$

By the definition of the infimum, such a sequence must exist.

Because X is compact, there is a *sub*-sequence $\{\underline{x}_{k_m}\}$ that converges to some $\underline{x}' \in X$ as $m \to \infty$, i.e.,

$$\lim_{m\to\infty} ||\underline{x}_{k_m} - \underline{x}'|| = 0. \tag{174}$$

We will show that $f(\underline{x}') = f^*$.

Because f is continuous, it must be that $f(\underline{x}_{k_m})$ converges to $f(\underline{x}')$, which is to say that for any $\varepsilon > 0$, $|f(\underline{x}_{k_m}) - f(\underline{x}')| < \varepsilon$ for all sufficiently large m. However, we thus have for all sufficiently large m,

$$0 \leqslant f(\underline{x}') - f^* \leqslant |f(\underline{x}') - f(\underline{x}_{k_m})| + f(\underline{x}_{k_m}) - f^* \leqslant \varepsilon + \frac{1}{k_m}. \tag{175}$$

Taking $m \to \infty$, the above gives that $0 \leqslant f(\underline{x}') - f^* \leqslant \varepsilon$. As this holds for $\varepsilon > 0$, taking $\varepsilon \to 0$ we have

$$f(\underline{x}') - f^* = 0, \tag{176}$$

or that $\underline{x}'$ realizes the minimal function value f^*, i.e., taking $\underline{x}^* = \underline{x}'$,

$$f(\underline{x}^*) = \min_{\underline{x}\in X} f(\underline{x}). \tag{177}$$

B Mathematical Background

Proposition 4 *Let A be an n by m matrix. If the columns of A are linearly independent, then $A^T A$ is positive definite.*

Proof. Note that for any $\underline{x} \in \mathbb{R}^m$, we have $\underline{x}^{\mathrm{T}} A^{\mathrm{T}} A \underline{x} = (A\underline{x})^{\mathrm{T}}(A\underline{x}) = ||A\underline{x}||^2$. If $\underline{x}^{\mathrm{T}} A^{\mathrm{T}} A \underline{x} = 0$, we therefore have that $A\underline{x} = 0$. However, if the columns of A are linearly independent, then the only solution is $\underline{x} = 0$. Hence, $A^{\mathrm{T}} A$ is positive definite. $\square$