CISC 371 Class 19

Constrained Optimization Problems

Texts: [1] pp. 147–155; [2] pp. 127–138

Main Concepts:
- *Optimization with equality constraints*
- *Optimization with inequality constraints*
- *Convex optimization problems*

Sample Problem, Machine Inference: where, inside a circle, is a hyperplane decision minimized?

We will now turn our attention to the third kind of optimization problem that is commonly encountered, which is *constrained optimization*. This kind of problem is the minimization of an objective function subject to constraints.

Example: Solutions to a Linear Constraint

A constrained optimization problem can arise in contexts where the nature of the optimization is not immediately apparent. For example, suppose that we are presented with an equation that seems to have an infinite number of possible solutions. An example from basic linear algebra is

$$\begin{array}{rrcl} & -w_1 + 2w_2 & = & 4 \\ \equiv & -w_1 + 2w_2 - 4 & = & 0 \end{array} \tag{19.1}$$

It is easy to verify that one solution to Equation 19.1 is $\vec{w}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

It is also easy to verify that there is an infinite set of solutions, of the form

$$\vec{l}(t) = \vec{w}_1 + t \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{19.2}$$

Equation 19.2 can be used to verify that $\vec{l}(-2)$ is another solution to Equation 19.1. The solutions, which are a line in 2D, are plotted in Figure 19.1(A).

Next, suppose that we are asked to select a solution $\vec{w}^*$ Equation 19.1 that has the minimum squared distance to the point $\vec{w}_0 = \begin{bmatrix} 1.5 \\ -1 \end{bmatrix}$. This is illustrated in Figure 19.1(B).

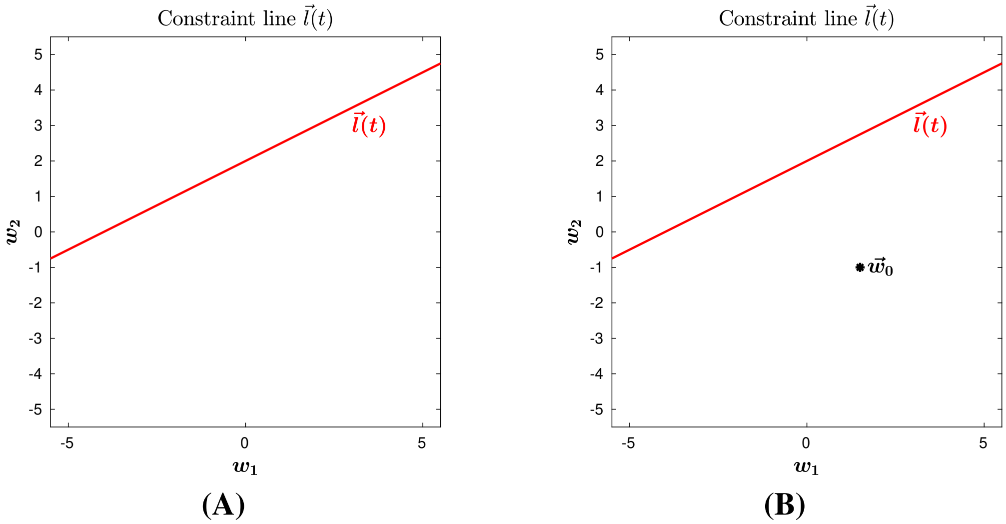


Figure 19.1: The set of solutions to Equation 19.1 are a line in $\mathbb{R}^2$. (A) The set, plotted as a black line. (B) A point $\vec{w}_0$, plotted in blue; a problem is to find the point of closest approach to $\vec{w}_0$ that is on the line $\vec{l}(t)$.

As stated, this is a problem of closest approach. We have a parametric equation in the scalar t, and a point $\vec{w}_0$; the problem is to find the point in the set $\vec{l}(t)$ that is closest to the point $\vec{w}_0$.

<u>**Problem:**</u> closest approach, squared length

$$
\begin{array}{rrcl}
\text{Find:} & & & \\
 & \vec{w}^* & = & \vec{l}(t^*) \\
\text{where:} & & & \\
 & t^* & = & \operatorname*{argmin}_{t \in \mathbb{R}} \|\vec{l}(t) - \vec{w}_0\|^2
\end{array}
\tag{19.3}
$$

To we find a vector $\vec{w}^*$ that satisfies both of these conditions, we will "invert" the requirements of Problem 19.3. Instead, we will pose this as an optimization problem – finding the minimum squared length of the vector $[\vec{w} - \vec{w}_0]$ – that has Equation 19.1 as a single linear equality constraint. In this course, we will use the following presentation for a constrained optimization problem, where the linear equality constraint is re-written as being equal to zero.

We will call the set of points that satisfy all of the given constraints the set of *feasible points*. This term is defined in the extra notes for this class.

Problem: squared length, linear equality

$$\begin{aligned} \text{Find:} \quad & \\ \vec{w}^* &= \underset{\vec{w}\in\mathbb{R}^2}{\operatorname{argmin}} [\vec{w}-\vec{w}_0]^T[\vec{w}-\vec{w}_0] \\ \text{where:} \quad & \\ \begin{bmatrix}-1 & 2\end{bmatrix}\vec{w} - 4 &= 0 \end{aligned} \tag{19.4}$$

Problem 19.4 is presented graphically in Figure 19.2. The linear constraint of Equation 19.1 is the black line and the contours of the squared-length function are shown in blue. The optimal squared length of the vector is shown in green.

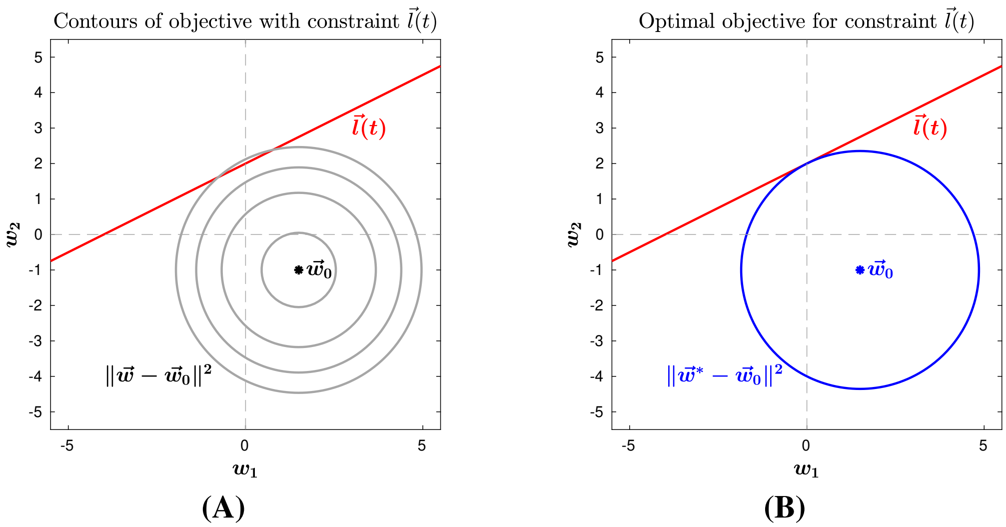


Figure 19.2: The set of solutions to Problem 19.4 are a line in $\mathbb{R}^2$, shown in black. (A) Contours of the squared length are shown in blue. (B) The contour of the optimal vector length is shown in green.

19.1 Equality Constraints and Inequality Constraints

We will refer to Problem 19.4 as an example of a quadratic problem with a single equality constraint. The problem is quadratic because the objective function is a second-order function of the independent vector variable $\vec{w}$. The constraint is linear because it has the form

$$\begin{aligned} M\vec{w} &= \vec{c} \\ \text{or} \quad M\vec{w}-\vec{c} &= \vec{0} \end{aligned}$$

A related linear constraint is an *inequality* constraint. An example is the above objective function – the squared distance of $\vec{w}$ to the point $\vec{w}_0$ – plus the constraint that the optimal point $\vec{w}^*$ is in the half-plane

$$\begin{aligned} w_1 - w_2 &\leq -4 \\ \text{or} \quad w_1 - w_2 + 4 &\leq 0 \end{aligned} \tag{19.5}$$

We will formally state this problem as

Problem: squared length, linear inequality

$$\begin{aligned} \text{Find:} \quad & \\ \vec{w}^* &= \underset{\vec{w} \in \mathbb{R}^2}{\operatorname{argmin}} [\vec{w} - \vec{w}_0]^T [\vec{w} - \vec{w}_0] \\ \text{where:} \quad & \\ \begin{bmatrix} 1 & -1 \end{bmatrix} \vec{w} + 4 &\leq 0 \end{aligned} \tag{19.6}$$

The inequality constraint of Problem 19.6 is shown in Figure 19.3(A) as a shaded region. The contours of the squared-length function are shown in Figure 19.3(B) in blue.

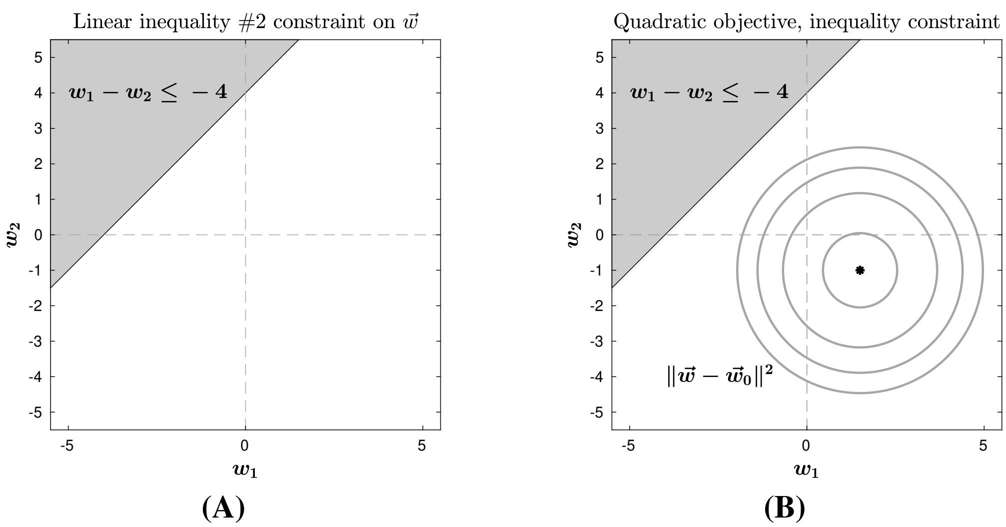


Figure 19.3: The set of points $\mathbb{F}$ that are feasible solutions to Problem 19.6 are a half-space. (A) The level set of the linear inequality constraint is shown as a shaded half-space. (B) Contours of the objective function are shown in blue.

The objective function for constrained optimization might be an affine function, subject to linear inequality constraints. For example, we might use the left-hand side of Equation 19.1 as the objective, where we seek the a vector along a line. We might then use the linear inequality constraint

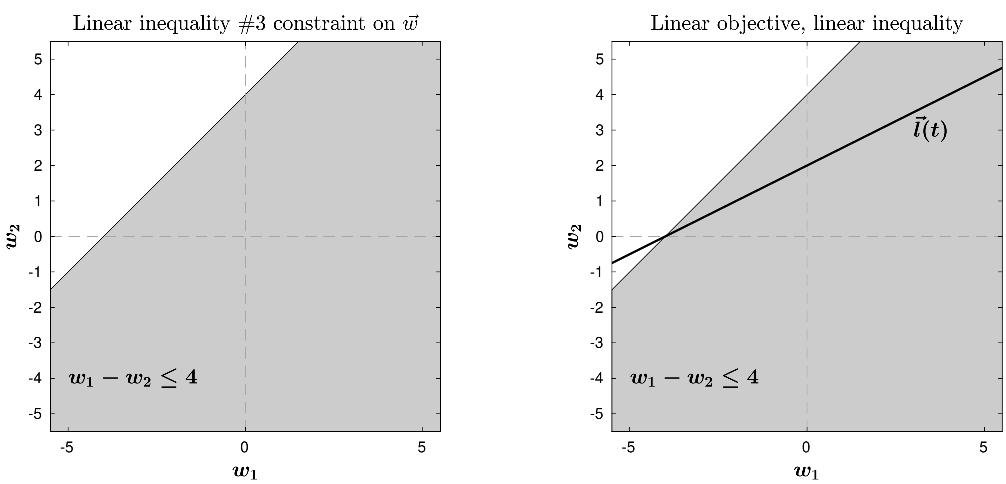
$$\begin{aligned} -w_1 + w_2 &\leq 4 \\ \text{or} \quad -w_1 + w_2 - 4 &\leq 0 \end{aligned} \tag{19.7}$$

Combining the left-hand side of Equation 19.1 with Equation 19.7 would give us an affine objective function with an linear inequality constraint that we could formalize as

Problem: linear objective length, linear inequality

$$\begin{aligned} &\text{Find:} \\ &\qquad \vec{w}^* = \underset{\vec{w}\in\mathbb{R}^2}{\operatorname{argmin}} -w_1 + 2w_2 - 4 \\ &\text{where:} \\ &\qquad \begin{bmatrix}-1 & 1\end{bmatrix} \vec{w} - 4 \leq 0 \end{aligned} \tag{19.8}$$

The inequality constraint of Problem 19.8 is shown in Figure 19.4(A) as a shaded region. The affine objective function is shown in Figure 19.4(B) in blue. Graphically, the solution is the "leftmost" point on the line $\vec{l}(t)$ that lies within the shaded region, which is the set of points $\mathbb{F}$ that are feasible solutions to Problem 19.8.



(A) **(B)**

Figure 19.4: The set of points $\mathbb{F}$ that are feasible solutions to Problem 19.8 are a half-space. (A) The level set of the linear inequality constraint is shown as a shaded half-space. (B) The objective function is shown in blue.

19.2 Extra Notes for Constrained Optimization Problems

The three example problems in this class are examples of more general problems in optimization.

Entire textbooks, and courses at many educaitonal levels, are devoted to *convex* optimization problems. These are problems in which the objective function is a convex function. In addition, there may be linear equality constraints, or inequality constraints that are have solutions restricted to convex sets.

When we use a convex set, we will define the set as the zero level set of a convex function. We will call such a function a *property* function.

<u>Definition:</u> property function in convex optimization

For any $i \in \mathbb{R}_+$, that a function $p_i : \mathbb{R}^n \to \mathbb{R}$ is a *property* function is defined as:

$$p_i \text{ is convex}$$

The set of points that is the intersection of zero level sets of all of the property functions is the *feasible set* for the optimization problem.

<u>Definition:</u> feasible point in convex optimization

For any m property functions each of the form $p_i : \mathbb{R}^n \to \mathbb{R}$, that a point $\vec{w} \in \mathbb{R}^n$ is a *feasible point* is defined as:

$$(i \in \mathbb{R}_+) \wedge (i \leq m) \to (p_i(\vec{w}) \leq 0)$$

<u>Definition:</u> feasible set $\mathbb{F}$ of an optimization problem

For a constrained optimization problem, the *feasible set* $\mathbb{F}$ is defined as:

$$\{\vec{w} : \vec{w} \text{ is a feasible point}\}$$

The set of solutions to a linear equation $M\vec{w} = \vec{c}$ is an important convex set. Recall, from Theorem 18.17, that an affine map is a convex function. The level set of a convex function is a convex set, so any level set of an affine map is a convex set.

Theorem: convexity of the zero level set of an affine map

For any positive integer $l < n$, any matrix $M_{l \times n}$ for which $\text{rank}(M) = l$, and any vector $\vec{c} \in \mathbb{R}^l$,

$$\textit{The set } \{\vec{w} : (\vec{w} \in \mathbb{R}^n) \wedge (M\vec{w} - \vec{c} \leq \vec{0})\} \textit{ is convex} \tag{19.9}$$

Proof: Let $\vec{u} \in \mathbb{R}^n$ such that $M\vec{u} - \vec{c} \leq \vec{0}$, $\vec{v} \in \mathbb{R}^n$ such that $M\vec{v} - \vec{c} \leq \vec{0}$, and $\theta \in \mathbb{R}_+$ such that $0 \leq \theta \leq 1$. Then

$$\begin{aligned}
(1-\theta)[M\vec{u} - \vec{c}] &\leq (1-\theta)[\vec{0}] \\
\theta[M\vec{v} - \vec{c}] &\leq \theta[\vec{0}]
\end{aligned}$$

and

$$\begin{aligned}
M[(1-\theta)\vec{u} + \theta\vec{v}] - \vec{c} &= M[(1-\theta)\vec{u} + \theta\vec{v}] - [(1-\theta)\vec{c} + \theta\vec{c}] \\
&= [(1-\theta)M\vec{u} - (1-\theta)\vec{c}] + [\theta M\vec{v} - \theta\vec{c}] \\
&= (1-\theta)[M\vec{u} - \vec{c}] + \theta[M\vec{v} - \vec{c}] \\
\text{so} \quad M[\theta\vec{u} + (1-\theta)\vec{v}] - \vec{c} &\leq (1-\theta)[\vec{0}] + \theta[\vec{0}] \\
&\leq \vec{0}
\end{aligned}$$

Observation: We can directly conclude that the boundary of the level set, which is the set of vectors that solves $M\vec{w} = \vec{c}$, is a convex set.

19.2.1 Convex Optimization

We can use these results to define a powerful class of optimization problems.

Definition: standard form of constrained convex optimization

For any convex function $f : \mathbb{R}^n \to \mathbb{R}$, any m property functions each of the form $p_i : \mathbb{R}^n \to \mathbb{R}$, any positive integer $l < n$, any matrix $M_{l \times n}$ for which $\text{rank}(M) = l$, and any vector $\vec{c} \in \mathbb{R}^l$, a *convex problem in standard form* is defined as:

$$\begin{aligned}
\vec{w}^* &= \underset{\vec{w} \in \mathbb{R}^n}{\text{argmin}}\, f(\vec{w}) \\
\text{where:} & \\
\forall_{i \leq m} p_i(\vec{w}) &\leq 0 \\
M\vec{w} - \vec{c} &= \vec{0}
\end{aligned} \tag{19.10}$$

19.2.2 Quadratic Programming

For historical reasons, *convex programming* is a name for a convex optimization problem in which:

- The objective function $f(\vec{w})$ is a quadratic function
$$f(\vec{w}) = \vec{w}^T K \vec{w} + \vec{q}^T \vec{w}$$
- Each property function, if it exists, is a linear function

We will use a simple definition for quadratic programming.

Definition: constrained quadratic optimization

For any symmetric positive semidefinite matrix $K \succeq 0$, any vector $\vec{q} \in \mathbb{R}^n$, any integer m, any matrix $A_{m\times n}$, any vector $\vec{b} \in \mathbb{R}^m$, any positive integer $l < n$, any matrix $M_{l\times n}$ for which $\text{rank}(M) = l$, and any vector $\vec{c} \in \mathbb{R}^l$, a ***constrained quadratic optimization problem*** is defined as:

$$\begin{aligned} \vec{w}^* &= \operatorname*{argmin}_{\vec{w}\in\mathbb{R}^n} f(\vec{w}) \\ \text{where:} & \\ f(\vec{w}) &= \vec{w}^T K \vec{w} + \vec{q}^T \vec{w} \\ A\vec{w} - \vec{b} &\leq \vec{0} \\ M\vec{w} - \vec{c} &= \vec{0} \end{aligned} \tag{19.11}$$

The first two examples in this class are constrained quadratic optimization problems.

19.2.3 Linear Programming

For historical reasons, *linear programming* is optimization of a linear objective function with linear constraints. The standard form is usually represented as a maximization – rather than as a minimization – and the independent variables in the vector argument are standardized by linear or affine transformations.

Definition: linear programming

For any vector $\vec{q} \in \mathbb{R}^n$, any integer m, any matrix $A_{m \times n}$, and any vector $\vec{b} \in \mathbb{R}^m$, a *linear programming problem in standard form* is defined as:

$$\begin{array}{rrcl} & \vec{w}^* & = & \underset{\vec{w} \in \mathbb{R}^n}{\operatorname{argmax}} f(\vec{w}) \\ \text{where:} & & & \\ & f(\vec{w}) & = & \vec{q}^T \vec{w} \\ & A\vec{w} - \vec{b} & \leq & \vec{0} \\ & \vec{w} & \geq & \vec{0} \end{array} \tag{19.12}$$

The third example in this class is a linear programming problem in which the constraint on non-negativity of the feasible solutions is omitted.

End of Extra Notes

References

[1] Beck A: Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB. Siam Press, 2014

[2] Boyd S, Vandenberghe L: Convex Optimization. Cambridge University Press, 2004