Non-linear Programming and Solver

Vahan Mkrtchyan [1]

[1] Lane Department of Computer Science and Electrical Engineering
West Virginia University

February 17, 2015

Outline

1 Motivating Examples

Outline

1. Motivating Examples
2. Problem Formulation

Outline

1. Motivating Examples
2. Problem Formulation
3. Constrained Optimization

Outline

1. Motivating Examples
2. Problem Formulation
3. Constrained Optimization
4. The generalized reduced gradient method

Outline

1. Motivating Examples
2. Problem Formulation
3. Constrained Optimization
4. The generalized reduced gradient method
5. Non-smooth Optimization: Subgradient methods

Portfolio Optimization

Portfolio Optimization

Problem formulation

Portfolio Optimization

Problem formulation

Suppose we have a sum of money M to split among three managed investment funds, which claim to offer percentage rates of return r_1, r_2 and r_3.

Portfolio Optimization

Problem formulation

Suppose we have a sum of money M to split among three managed investment funds, which claim to offer percentage rates of return r_1, r_2 and r_3.

If we invest amounts y_1, y_2 and y_3, we can expect our overall return to be

$$R = \frac{r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3}{M}\%.$$

Portfolio Optimization

Problem formulation

Suppose we have a sum of money M to split among three managed investment funds, which claim to offer percentage rates of return r_1, r_2 and r_3.

If we invest amounts y_1, y_2 and y_3, we can expect our overall return to be

$$R = \frac{r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3}{M} \%.$$

Assume that the management charge associated with the i-th fund is calculated as $c_i \cdot y_i$.

Portfolio Optimization

Problem formulation

Suppose we have a sum of money M to split among three managed investment funds, which claim to offer percentage rates of return r_1, r_2 and r_3.

If we invest amounts y_1, y_2 and y_3, we can expect our overall return to be

$$R = \frac{r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3}{M} \%.$$

Assume that the management charge associated with the i-th fund is calculated as $c_i \cdot y_i$.

Then the total cost of investment is

$$C = c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3.$$

Portfolio Optimization

Problem formulation

Suppose we have a sum of money M to split among three managed investment funds, which claim to offer percentage rates of return r_1, r_2 and r_3.

If we invest amounts y_1, y_2 and y_3, we can expect our overall return to be

$$R = \frac{r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3}{M}\%.$$

Assume that the management charge associated with the i-th fund is calculated as $c_i \cdot y_i$.

Then the total cost of investment is

$$C = c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3.$$

Now, assume that we are aiming for a return $R_\rho\%$, and that we want to pay the least charges to achieve this.

Portfolio Optimization

Portfolio Optimization

Problem modeling

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3$$

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3$$
$$r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho,$$

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3$$
$$r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho,$$
$$y_1 + y_2 + y_3 = M$$

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

$$\begin{aligned} \min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3 \\ r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho, \\ y_1 + y_2 + y_3 = M \\ y_1 \geq 0, y_2 \geq 0, y_3 \geq 0. \end{aligned}$$

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3$$
$$r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho,$$
$$y_1 + y_2 + y_3 = M$$
$$y_1 \geq 0, y_2 \geq 0, y_3 \geq 0.$$

Remark

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3$$
$$r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho,$$
$$y_1 + y_2 + y_3 = M$$
$$y_1 \geq 0, y_2 \geq 0, y_3 \geq 0.$$

Remark

The last inequalities are included because investments must obviously be positive.

Portfolio Optimization

Problem modeling

The amounts y_1, y_2 and y_3 need to be chosen so that the following conditions are satisfied:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3$$
$$r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho,$$
$$y_1 + y_2 + y_3 = M$$
$$y_1 \geq 0, y_2 \geq 0, y_3 \geq 0.$$

Remark

The last inequalities are included because investments must obviously be positive.

If we tried to solve the problem without them, an optimization algorithm would attempt to reduce costs by making one or more of the y_i large and negative.

Portfolio Optimization

Portfolio Optimization

Problem formulation

Portfolio Optimization

Problem formulation

Now, assume that an attempt to make a negative investment would be penalized by a very high management charge.

Portfolio Optimization

Problem formulation

Now, assume that an attempt to make a negative investment would be penalized by a very high management charge.

Assume that the charge is given by the function $K \cdot \Psi(x)$, where

$$\Psi(x) = \begin{cases} x^2, & \text{if } x < 0 \\ 0, & \text{otherwise.} \end{cases}$$

and K is a large positive constant.

Portfolio Optimization

Problem formulation

Now, assume that an attempt to make a negative investment would be penalized by a very high management charge.

Assume that the charge is given by the function $K \cdot \Psi(x)$, where

$$\Psi(x) = \begin{cases} x^2, & \text{if } x < 0 \\ 0, & \text{otherwise.} \end{cases}$$

and K is a large positive constant.

Then the problem becomes:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3 + K \cdot \sum_{i=1}^{3} \Psi(y_i)$$

Portfolio Optimization

Problem formulation

Now, assume that an attempt to make a negative investment would be penalized by a very high management charge.

Assume that the charge is given by the function $K \cdot \Psi(x)$, where

$$\Psi(x) = \begin{cases} x^2, & \text{if } x < 0 \\ 0, & \text{otherwise.} \end{cases}$$

and K is a large positive constant.

Then the problem becomes:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3 + K \cdot \sum_{i=1}^{3} \Psi(y_i)$$

$$r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho,$$

Portfolio Optimization

Problem formulation

Now, assume that an attempt to make a negative investment would be penalized by a very high management charge.

Assume that the charge is given by the function $K \cdot \Psi(x)$, where

$$\Psi(x) = \begin{cases} x^2, & \text{if } x < 0 \\ 0, & \text{otherwise.} \end{cases}$$

and K is a large positive constant.

Then the problem becomes:

$$\min c_1 \cdot y_1 + c_2 \cdot y_2 + c_3 \cdot y_3 + K \cdot \sum_{i=1}^{3} \Psi(y_i)$$
$$r_1 \cdot y_1 + r_2 \cdot y_2 + r_3 \cdot y_3 = M \cdot R_\rho,$$
$$y_1 + y_2 + y_3 = M$$

Portfolio return and risk

Portfolio return and risk

Main concepts

Portfolio return and risk

Main concepts

Suppose we have a history of percentage returns, over m time periods, for each of a group of n assets (such as shares, bonds etc.).

Portfolio return and risk

Main concepts

Suppose we have a history of percentage returns, over m time periods, for each of a group of n assets (such as shares, bonds etc.).

We can use this information as a guide to future investments.

Portfolio return and risk

Main concepts

Suppose we have a history of percentage returns, over m time periods, for each of a group of n assets (such as shares, bonds etc.).

We can use this information as a guide to future investments.

As an example, consider the following data for three assets over six months.

Portfolio return and risk

Main concepts

Suppose we have a history of percentage returns, over *m* time periods, for each of a group of *n* assets (such as shares, bonds etc.).

We can use this information as a guide to future investments.

As an example, consider the following data for three assets over six months.

Monthly rates of return on three assets

Assets/Months	January	February	March	April	May	June
Assets 1	1.2	1.3	1.4	1.5	1.1	1.2
Assets 2	1.3	1.0	0.8	0.9	1.4	1.3
Assets 3	0.9	1.1	1.0	1.1	1.1	1.3

Portfolio return and risk

Portfolio return and risk

Main concepts

Portfolio return and risk

Main concepts

In general, we can calculate the mean return $\bar{r}_i$ for each asset as

$$\bar{r}_i = \frac{\sum_{j=1}^{m} r_{ij}}{m},$$

Portfolio return and risk

Main concepts

In general, we can calculate the mean return $\bar{r}_i$ for each asset as

$$\bar{r}_i = \frac{\sum_{j=1}^{m} r_{ij}}{m},$$

where r_{ij} denotes the return on asset i in period j.

Portfolio return and risk

Portfolio return and risk

Portfolio and expected Portfolio return

Portfolio return and risk

Portfolio and expected Portfolio return

If we spread an investment across the n assets and if y_i denotes the fraction invested in asset i then the values of the y_i define a portfolio.

Portfolio return and risk

Portfolio and expected Portfolio return

If we spread an investment across the n assets and if y_i denotes the fraction invested in asset i then the values of the y_i define a portfolio.

Since all investment must be split between the n assets, the invested fractions must satisfy

$$S = \sum_{i=1}^{n} y_i = 1$$

Portfolio return and risk

Portfolio and expected Portfolio return

If we spread an investment across the n assets and if y_i denotes the fraction invested in asset i then the values of the y_i define a portfolio.

Since all investment must be split between the n assets, the invested fractions must satisfy

$$S = \sum_{i=1}^{n} y_i = 1$$

The expected portfolio return is given by

$$R = \sum_{i=1}^{n} \bar{r}_i \cdot y_i.$$

Portfolio return and risk

Portfolio return and risk

Portfolio risk

Portfolio return and risk

Portfolio risk

The risk associated with a particular portfolio is determined from variances and covariances that can be calculated from the history of returns r_{ij}.

Portfolio return and risk

Portfolio risk

The risk associated with a particular portfolio is determined from variances and covariances that can be calculated from the history of returns r_{ij}.

The variance of asset i is

$$\sigma_i^2 = \frac{\sum_{j=1}^{m}(r_{ij} - \bar{r}_i)^2}{m},$$

Portfolio return and risk

Portfolio risk

The risk associated with a particular portfolio is determined from variances and covariances that can be calculated from the history of returns r_{ij}.

The variance of asset i is

$$\sigma_i^2 = \frac{\sum_{j=1}^{m} (r_{ij} - \bar{r}_i)^2}{m},$$

while the covariance of assets i and k is

$$\sigma_{ik} = \frac{\sum_{j=1}^{m} (r_{ij} - \bar{r}_i) \cdot (r_{kj} - \bar{r}_k)}{m}.$$

Portfolio return and risk

Portfolio return and risk

Portfolio variance

Portfolio return and risk

Portfolio variance

The variance of the portfolio defined by the investment fractions y_1, ...,y_n is

$$V = \sum_{i=1}^{n} \sigma_i^2 \cdot y_i^2 + 2 \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sigma_{ij} \cdot y_i \cdot y_j,$$

which can be used as a measure of portfolio risk.

Portfolio return and risk: simple notation

Portfolio return and risk: simple notation

Matrix-vector notation

Portfolio return and risk: simple notation

Matrix-vector notation

The return and risk functions can be written more conveniently using matrix-vector notation.

Portfolio return and risk: simple notation

Matrix-vector notation

The return and risk functions can be written more conveniently using matrix-vector notation.

The expression for expected return can be written as:

$$R = \bar{\mathbf{r}} \cdot \mathbf{y},$$

Portfolio return and risk: simple notation

Matrix-vector notation

The return and risk functions can be written more conveniently using matrix-vector notation.

The expression for expected return can be written as:

$$R = \bar{\mathbf{r}} \cdot \mathbf{y},$$

and

$$V = \mathbf{y}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{y},$$

Portfolio return and risk: simple notation

Matrix-vector notation

The return and risk functions can be written more conveniently using matrix-vector notation.

The expression for expected return can be written as:

$$R = \bar{\mathbf{r}} \cdot \mathbf{y},$$

and

$$V = \mathbf{y}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{y},$$

where $\bar{\mathbf{r}}$ denotes the vector of mean returns $\bar{r}_i$, and $\mathbf{Q} = \|\sigma_{ij}\|$.

Portfolio return and risk: simple notation

Matrix-vector notation

The return and risk functions can be written more conveniently using matrix-vector notation.

The expression for expected return can be written as:

$$R = \bar{\mathbf{r}} \cdot \mathbf{y},$$

and

$$V = \mathbf{y}^{\mathbf{T}} \cdot \mathbf{Q} \cdot \mathbf{y},$$

where $\bar{\mathbf{r}}$ denotes the vector of mean returns $\bar{r}_i$, and $\mathbf{Q} = \|\sigma_{ij}\|$.

The constraint on partitions can be written as

$$S = \mathbf{e} \cdot \mathbf{y},$$

where $\mathbf{e} = (1, 1, ..., 1)$.

The basic minimum risk problem

The basic minimum risk problem

Problem formulation

The basic minimum risk problem

Problem formulation

A major concern in portfolio selection is the minimization of risk.

The basic minimum risk problem

Problem formulation

A major concern in portfolio selection is the minimization of risk.

In its simplest form, this means finding invested fractions $y_1, ..., y_n$, to solve the problem

The basic minimum risk problem

Problem formulation

A major concern in portfolio selection is the minimization of risk.

In its simplest form, this means finding invested fractions $y_1, \ldots, y_n$, to solve the problem

$$\min \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}$$

The basic minimum risk problem

Problem formulation

A major concern in portfolio selection is the minimization of risk.

In its simplest form, this means finding invested fractions $y_1, ..., y_n$, to solve the problem

$$\min \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}$$
$$\mathbf{e} \cdot \mathbf{y} = 1,$$

Optimizing return and risk

Optimizing return and risk

Problem formulation

Optimizing return and risk

Problem formulation

The solution to the basic problem can sometimes be useful.

Optimizing return and risk

Problem formulation

The solution to the basic problem can sometimes be useful.

But in practice we will normally be interested in both risk and return rather than risk on its own.

Optimizing return and risk

Problem formulation

The solution to the basic problem can sometimes be useful.

But in practice we will normally be interested in both risk and return rather than risk on its own.

In a rather general way, we can say that an optimal portfolio is one which gives ”biggest return at lowest risk”.

Optimizing return and risk

Problem formulation

The solution to the basic problem can sometimes be useful.

But in practice we will normally be interested in both risk and return rather than risk on its own.

In a rather general way, we can say that an optimal portfolio is one which gives "biggest return at lowest risk".

One way of trying to determine such a portfolio is to consider a composite function such as

$$F = -R + \rho \cdot V = -\bar{\mathbf{r}} \cdot \mathbf{y} + \rho \cdot \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}.$$

Optimizing return and risk

Problem formulation

The solution to the basic problem can sometimes be useful.

But in practice we will normally be interested in both risk and return rather than risk on its own.

In a rather general way, we can say that an optimal portfolio is one which gives ”biggest return at lowest risk”.

One way of trying to determine such a portfolio is to consider a composite function such as

$$F = -R + \rho \cdot V = -\bar{\mathbf{r}} \cdot \mathbf{y} + \rho \cdot \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}.$$

The first term is the negative of the expected return and the second term is a multiple of the risk.

Optimizing return and risk

Problem formulation

The solution to the basic problem can sometimes be useful.

But in practice we will normally be interested in both risk and return rather than risk on its own.

In a rather general way, we can say that an optimal portfolio is one which gives ”biggest return at lowest risk”.

One way of trying to determine such a portfolio is to consider a composite function such as

$$F = -R + \rho \cdot V = -\bar{\mathbf{r}} \cdot \mathbf{y} + \rho \cdot \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}.$$

The first term is the negative of the expected return and the second term is a multiple of the risk.

If we choose invested fractions y_i to minimize F, we can expect to obtain a large value for return coupled with a small value for risk.

Optimizing return and risk

Problem formulation

The solution to the basic problem can sometimes be useful.

But in practice we will normally be interested in both risk and return rather than risk on its own.

In a rather general way, we can say that an optimal portfolio is one which gives ”biggest return at lowest risk”.

One way of trying to determine such a portfolio is to consider a composite function such as

$$F = -R + \rho \cdot V = -\bar{\mathbf{r}} \cdot \mathbf{y} + \rho \cdot \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}.$$

The first term is the negative of the expected return and the second term is a multiple of the risk.

If we choose invested fractions y_i to minimize F, we can expect to obtain a large value for return coupled with a small value for risk.

The positive constant ρ controls the balance between return and risk.

Optimizing return and risk

Mathematical formulation

Optimizing return and risk

Mathematical formulation

$$\min -\bar{\mathbf{r}} \cdot \mathbf{y} + \rho \cdot \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}$$

Optimizing return and risk

Mathematical formulation

$$\min -\bar{\mathbf{r}} \cdot \mathbf{y} + \rho \cdot \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}$$
$$\mathbf{e} \cdot \mathbf{y} = 1,$$

Minimum risk for specified return

Minimum risk for specified return

Problem formulation

Minimum risk for specified return

Problem formulation

The previous problem allows us to balance risk and return according to the choice of the parameter ρ.

Minimum risk for specified return

Problem formulation

The previous problem allows us to balance risk and return according to the choice of the parameter ρ.

Another approach could be to fix a target value for return, say $R_\rho\%$, and to consider the problem

Minimum risk for specified return

Problem formulation

The previous problem allows us to balance risk and return according to the choice of the parameter ρ.

Another approach could be to fix a target value for return, say $R_\rho\%$, and to consider the problem

$$\min \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}$$

Minimum risk for specified return

Problem formulation

The previous problem allows us to balance risk and return according to the choice of the parameter ρ.

Another approach could be to fix a target value for return, say $R_\rho\%$, and to consider the problem

$$\min \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y}$$
$$\bar{\mathbf{r}} \cdot \mathbf{y} = R_\rho,$$

Minimum risk for specified return

Problem formulation

The previous problem allows us to balance risk and return according to the choice of the parameter ρ.

Another approach could be to fix a target value for return, say $R_\rho\%$, and to consider the problem

$$\begin{aligned} \min \mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y} \\ \bar{\mathbf{r}} \cdot \mathbf{y} = R_\rho, \\ \mathbf{e} \cdot \mathbf{y} = 1. \end{aligned}$$

Maximum return problem

Maximum return problem

Problem formulation

Maximum return problem

Problem formulation

Suppose we want to fix an acceptable level of risk (as V_a, say) and then to maximize the expected return.

Maximum return problem

Problem formulation

Suppose we want to fix an acceptable level of risk (as V_a, say) and then to maximize the expected return.

This can be posed as the constrained minimization problem

$$\max \bar{\mathbf{r}} \cdot \mathbf{y}$$

Maximum return problem

Problem formulation

Suppose we want to fix an acceptable level of risk (as V_a, say) and then to maximize the expected return.

This can be posed as the constrained minimization problem

$$\max \bar{\mathbf{r}} \cdot \mathbf{y}$$

$$\mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y} = V_a,$$

Maximum return problem

Problem formulation

Suppose we want to fix an acceptable level of risk (as V_a, say) and then to maximize the expected return.

This can be posed as the constrained minimization problem

$$\max \bar{\mathbf{r}} \cdot \mathbf{y}$$
$$\mathbf{y}^T \cdot \mathbf{Q} \cdot \mathbf{y} = V_a,$$
$$\mathbf{e} \cdot \mathbf{y} = 1.$$

Optimization Theory

Optimization Theory

Fundamentals

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility,

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness,

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness, Discrete).

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness, Discrete).
- Global minimizer (strict).

Optimization Theory

Fundamentals

Given a function $f : R^n \to R$ and a set $S \subseteq R^n$, the problem of finding an $\mathbf{x}^* \in R^n$ that solves

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\mathbf{x} \in S$$

is called an optimization problem.

Features of an optimization problem

- Decision variables.
- Objective function.
- Feasible region (Infeasibility, Unboundedness, Discrete).
- Global minimizer (strict).
- Local minimizer.

Non-linear programming

Non-linear programming

Problem Formulation

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programs

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programs

Such optimization problems usually are called *non-linear programming problems*

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programs

Such optimization problems usually are called *non-linear programming problems* or *non-linear programs.*

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programs

Such optimization problems usually are called *non-linear programming problems* or *non-linear programs.*

Constrained or Unconstrained Optimization

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programs

Such optimization problems usually are called *non-linear programming problems* or *non-linear programs.*

Constrained or Unconstrained Optimization

A non-linear program in which the set $\mathcal{E} \cup \mathcal{I}$ is empty, is called an unconstrained program.

Non-linear programming

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programs

Such optimization problems usually are called *non-linear programming problems* or *non-linear programs.*

Constrained or Unconstrained Optimization

A non-linear program in which the set $\mathcal{E} \cup \mathcal{I}$ is empty, is called an unconstrained program.

Otherwise, it is constrained.

Non-linear programs arise

Probabilistic elements

Non-linear programs arise

Probabilistic elements

Nonlinearities may arise when some of the coefficients in the model are random variables.

Non-linear programs arise

Probabilistic elements

Nonlinearities may arise when some of the coefficients in the model are random variables.

For example, consider a linear program, where the right-hand sides are random.

Non-linear programs arise

Probabilistic elements

Nonlinearities may arise when some of the coefficients in the model are random variables.

For example, consider a linear program, where the right-hand sides are random.

We consider the case when we have two constraints.

Probabilistic elements in linear programming

Probabilistic elements in linear programming

Problem Formulation

Probabilistic elements in linear programming

Problem Formulation

$$\max c_1 \cdot x_1 + \ldots + c_n \cdot x_n$$

Probabilistic elements in linear programming

Problem Formulation

$$\max c_1 \cdot x_1 + \ldots + c_n \cdot x_n$$

$$a_{11} \cdot x_1 + \ldots + a_{1n} \cdot x_n \leq b_1$$

Probabilistic elements in linear programming

Problem Formulation

$$\max c_1 \cdot x_1 + ... + c_n \cdot x_n$$
$$a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n \leq b_1$$
$$a_{21} \cdot x_1 + ... + a_{2n} \cdot x_n \leq b_2$$

Probabilistic elements in linear programming

Problem Formulation

$$\max c_1 \cdot x_1 + \ldots + c_n \cdot x_n$$
$$a_{11} \cdot x_1 + \ldots + a_{1n} \cdot x_n \leq b_1$$
$$a_{21} \cdot x_1 + \ldots + a_{2n} \cdot x_n \leq b_2$$

where the coefficients b_1 and b_2 are independently distributed and $G_i(y)$ represents the probability that the random variable b_i is at least as large as y.

Probabilistic elements in linear programming

Probabilistic elements in linear programming

Problem Formulation

Probabilistic elements in linear programming

Problem Formulation

Assume that we would like to choose the variables $x_1, \ldots, x_n$, so that the joint probability of both of the constraints being satisfied is at least β.

Probabilistic elements in linear programming

Problem Formulation

Assume that we would like to choose the variables $x_1, ..., x_n$, so that the joint probability of both of the constraints being satisfied is at least β.

The last constraint mathematically can be written as follows:

$$Pr[a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n \leq b_1] \times Pr[a_{21} \cdot x_1 + ... + a_{2n} \cdot x_n \leq b_2] \geq \beta.$$

Probabilistic elements in linear programming

Problem Formulation

Assume that we would like to choose the variables $x_1, ..., x_n$, so that the joint probability of both of the constraints being satisfied is at least β.

The last constraint mathematically can be written as follows:

$$Pr[a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n \leq b_1] \times Pr[a_{21} \cdot x_1 + ... + a_{2n} \cdot x_n \leq b_2] \geq \beta.$$

Then this condition can be written as the following set of constraints:

$$-y_1 + a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n = 0$$

Probabilistic elements in linear programming

Problem Formulation

Assume that we would like to choose the variables $x_1, ..., x_n$, so that the joint probability of both of the constraints being satisfied is at least β.

The last constraint mathematically can be written as follows:

$$Pr[a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n \leq b_1] \times Pr[a_{21} \cdot x_1 + ... + a_{2n} \cdot x_n \leq b_2] \geq \beta.$$

Then this condition can be written as the following set of constraints:

$$-y_1 + a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n = 0$$
$$-y_2 + a_{21} \cdot x_1 + ... + a_{2n} \cdot x_n = 0$$

Probabilistic elements in linear programming

Problem Formulation

Assume that we would like to choose the variables $x_1, ..., x_n$, so that the joint probability of both of the constraints being satisfied is at least β.

The last constraint mathematically can be written as follows:

$$Pr[a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n \leq b_1] \times Pr[a_{21} \cdot x_1 + ... + a_{2n} \cdot x_n \leq b_2] \geq \beta.$$

Then this condition can be written as the following set of constraints:

$$-y_1 + a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n = 0$$
$$-y_2 + a_{21} \cdot x_1 + ... + a_{2n} \cdot x_n = 0$$
$$G_1(y_1) \times G_2(y_2) \geq \beta.$$

Constrained Optimization

Constrained Optimization

Problem Formulation

Constrained Optimization

Problem Formulation

$$\min f(\mathbf{x})$$

Constrained Optimization

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$

Constrained Optimization

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Constrained Optimization

Problem Formulation

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Here we assume that we have at least one constraint, i.e., the set $\mathcal{E} \cup \mathcal{I}$ is not empty.

Constrained Optimization

Problem Formulation

$$\begin{aligned} &\min f(\mathbf{x}) \\ g_i(\mathbf{x}) &= 0, i \in \mathcal{E} \\ g_i(\mathbf{x}) &\geq 0, i \in \mathcal{I} \end{aligned}$$

Here we assume that we have at least one constraint, i.e., the set $\mathcal{E} \cup \mathcal{I}$ is not empty.

Moreover, we assume that the functions f and g_i are continuously differentiable.

Constrained Optimization

Constrained Optimization

Lagrangian function of the problem

Constrained Optimization

Lagrangian function of the problem

The Lagrangian function (or Lagrangian) is defined as follows:

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot g_i(\mathbf{x}).$$

Constrained Optimization

Lagrangian function of the problem

The Lagrangian function (or Lagrangian) is defined as follows:

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot g_i(\mathbf{x}).$$

Why Lagrangian function?

Constrained Optimization

Lagrangian function of the problem

The Lagrangian function (or Lagrangian) is defined as follows:

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot g_i(\mathbf{x}).$$

Why Lagrangian function?

It turns out that for suitably chosen values of λ_i, minimizing the unconstrained Lagrangian function $L(\mathbf{x}, \lambda)$ is equivalent to solving the above constrained non-linear program.

Some definitions

Some definitions

Definition

Some definitions

Definition

A point $\mathbf{x}$ satisfying $g_i(\mathbf{x}) = 0, i \in \mathcal{E}$ and $g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$ is called a feasible solution to the non-linear program.

Some definitions

Definition

A point **x** satisfying $g_i(\mathbf{x}) = 0, i \in \mathcal{E}$ and $g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$ is called a feasible solution to the non-linear program.

Definition

Let **x** be a feasible solution to the non-linear program,

Some definitions

Definition

A point $\mathbf{x}$ satisfying $g_i(\mathbf{x}) = 0, i \in \mathcal{E}$ and $g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$ is called a feasible solution to the non-linear program.

Definition

Let $\mathbf{x}$ be a feasible solution to the non-linear program, and let $\mathcal{J} \subseteq \mathcal{I}$ be the set of indices for which $g_i(\mathbf{x}) \geq 0$ is satisfied with equality.

Some definitions

Definition

A point **x** satisfying $g_i(\mathbf{x}) = 0, i \in \mathcal{E}$ and $g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$ is called a feasible solution to the non-linear program.

Definition

Let **x** be a feasible solution to the non-linear program, and let $\mathcal{J} \subseteq \mathcal{I}$ be the set of indices for which $g_i(\mathbf{x}) \geq 0$ is satisfied with equality.

Then **x** is a regular point of the program, if the gradient vectors $\triangledown \mathbf{g_i}(\mathbf{x})$ for $i \in \mathcal{E} \cup \mathcal{J}$ are linearly independent.

Some examples

Some examples

Problem Formulation: Regular points

Some examples

Problem Formulation: Regular points

$$\max(x^2 + y^2)$$

Some examples

Problem Formulation: Regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$

Some examples

Problem Formulation: Regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$

Some examples

Problem Formulation: Regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$
$$x + y \leq 1$$

Some examples

Problem Formulation: Regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$
$$x + y \leq 1$$

In this example any feasible point is regular, since the gradients of the constraints are $(1, 0)$, $(0, 1)$ and $(-1, -1)$.

Some examples

Some examples

Problem Formulation: Non-regular points

Some examples

Problem Formulation: Non-regular points

$$\max(x^2 + y^2)$$

Some examples

Problem Formulation: Non-regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$

Some examples

Problem Formulation: Non-regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$

Some examples

Problem Formulation: Non-regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$
$$x^3 + 3 \cdot y^2 \leq 0$$

Some examples

Problem Formulation: Non-regular points

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$
$$x^3 + 3 \cdot y^2 \leq 0$$

In this example the feasible solution $(0, 0)$ is not a regular point since the gradients of the constraints are $(1, 0)$, $(0, 1)$ and $(0, 0)$.

First order necessary conditions

First order necessary conditions

KKT conditions

First order necessary conditions

KKT conditions

The conditions that will be presented in the upcoming three theorems are called Karush-Kuhn-Tucker (KKT) conditions after their inventors.

First order necessary conditions

KKT conditions

The conditions that will be presented in the upcoming three theorems are called Karush-Kuhn-Tucker (KKT) conditions after their inventors.

Theorem

First order necessary conditions

KKT conditions

The conditions that will be presented in the upcoming three theorems are called Karush-Kuhn-Tucker (KKT) conditions after their inventors.

Theorem

Let $\mathbf{x}^*$ *be a local minimizer of the non-linear problem,*

First order necessary conditions

KKT conditions

The conditions that will be presented in the upcoming three theorems are called Karush-Kuhn-Tucker (KKT) conditions after their inventors.

Theorem

Let $\mathbf{x}^*$ *be a local minimizer of the non-linear problem, and assume that* $\mathbf{x}^*$ *is a regular point for the constraints of the problem.*

First order necessary conditions

KKT conditions

The conditions that will be presented in the upcoming three theorems are called Karush-Kuhn-Tucker (KKT) conditions after their inventors.

Theorem

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Then there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ such that

$$\triangledown f(x^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot \triangledown g_i(x^*) = 0,$$

First order necessary conditions

KKT conditions

The conditions that will be presented in the upcoming three theorems are called Karush-Kuhn-Tucker (KKT) conditions after their inventors.

Theorem

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Then there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ such that

$$\bigtriangledown f(x^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot \bigtriangledown g_i(x^*) = 0,$$

$$\lambda_i \geq 0, i \in \mathcal{I},$$

First order necessary conditions

KKT conditions

The conditions that will be presented in the upcoming three theorems are called Karush-Kuhn-Tucker (KKT) conditions after their inventors.

Theorem

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Then there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ such that

$$\begin{aligned} \triangledown f(x^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot \triangledown g_i(x^*) = 0, \\ \lambda_i \geq 0, i \in \mathcal{I}, \\ \lambda_i \cdot g_i(\mathbf{x}^*) = 0, i \in \mathcal{I}. \end{aligned}$$

First order necessary conditions

First order necessary conditions

Problem Formulation: Example

First order necessary conditions

Problem Formulation: Example

$$\max(x^2 + y^2)$$

First order necessary conditions

Problem Formulation: Example

$$\max(x^2 + y^2)$$
$$x \geq 0$$

First order necessary conditions

Problem Formulation: Example

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$

First order necessary conditions

Problem Formulation: Example

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$
$$x + y \leq 1$$

First order necessary conditions

Problem Formulation: Example

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$
$$x + y \leq 1$$

Lagrangian function

First order necessary conditions

Problem Formulation: Example

$$\max(x^2 + y^2)$$
$$x \geq 0$$
$$y \geq 0$$
$$x + y \leq 1$$

Lagrangian function

$$L(x, y, \lambda) = -(x^2 + y^2) - \lambda_1 \cdot x - \lambda_2 \cdot y - \lambda_3 \cdot (1 - x - y).$$

First order necessary conditions

First order necessary conditions

The constraints

First order necessary conditions

The constraints

$$-2 \cdot x - \lambda_1 + \lambda_3 = 0$$

First order necessary conditions

The constraints

$$-2 \cdot x - \lambda_1 + \lambda_3 = 0$$
$$-2 \cdot y - \lambda_2 + \lambda_3 = 0$$

First order necessary conditions

The constraints

$$
\begin{aligned}
-2 \cdot x - \lambda_1 + \lambda_3 &= 0 \\
-2 \cdot y - \lambda_2 + \lambda_3 &= 0 \\
\lambda_1, \lambda_2, \lambda_3 &\geq 0
\end{aligned}
$$

First order necessary conditions

The constraints

$$
\begin{aligned}
-2 \cdot x - \lambda_1 + \lambda_3 &= 0 \\
-2 \cdot y - \lambda_2 + \lambda_3 &= 0 \\
\lambda_1, \lambda_2, \lambda_3 &\geq 0 \\
\lambda_1 \cdot x &= 0
\end{aligned}
$$

First order necessary conditions

The constraints

$$
\begin{aligned}
-2 \cdot x - \lambda_1 + \lambda_3 &= 0 \\
-2 \cdot y - \lambda_2 + \lambda_3 &= 0 \\
\lambda_1, \lambda_2, \lambda_3 &\geq 0 \\
\lambda_1 \cdot x &= 0 \\
\lambda_2 \cdot y &= 0
\end{aligned}
$$

First order necessary conditions

The constraints

$$
\begin{aligned}
-2 \cdot x - \lambda_1 + \lambda_3 &= 0 \\
-2 \cdot y - \lambda_2 + \lambda_3 &= 0 \\
\lambda_1, \lambda_2, \lambda_3 &\geq 0 \\
\lambda_1 \cdot x &= 0 \\
\lambda_2 \cdot y &= 0 \\
\lambda_3 \cdot (1 - x - y) &= 0
\end{aligned}
$$

First order necessary conditions

The constraints

$$
\begin{aligned}
-2 \cdot x - \lambda_1 + \lambda_3 &= 0 \\
-2 \cdot y - \lambda_2 + \lambda_3 &= 0 \\
\lambda_1, \lambda_2, \lambda_3 &\geq 0 \\
\lambda_1 \cdot x &= 0 \\
\lambda_2 \cdot y &= 0 \\
\lambda_3 \cdot (1 - x - y) &= 0
\end{aligned}
$$

Stationary points

First order necessary conditions

The constraints

$$-2 \cdot x - \lambda_1 + \lambda_3 = 0$$
$$-2 \cdot y - \lambda_2 + \lambda_3 = 0$$
$$\lambda_1, \lambda_2, \lambda_3 \geq 0$$
$$\lambda_1 \cdot x = 0$$
$$\lambda_2 \cdot y = 0$$
$$\lambda_3 \cdot (1 - x - y) = 0$$

Stationary points

If $\lambda_3 = 0$, it can be shown that $x = y = 0$.

First order necessary conditions

The constraints

$$
\begin{aligned}
-2 \cdot x - \lambda_1 + \lambda_3 &= 0 \\
-\,2 \cdot y - \lambda_2 + \lambda_3 &= 0 \\
\lambda_1, \lambda_2, \lambda_3 &\geq 0 \\
\lambda_1 \cdot x &= 0 \\
\lambda_2 \cdot y &= 0 \\
\lambda_3 \cdot (1 - x - y) &= 0
\end{aligned}
$$

Stationary points

If $\lambda_3 = 0$, it can be shown that $x = y = 0$.

On the other hand, if $\lambda_3 \neq 0$, then $x + y = 1$,

First order necessary conditions

The constraints

$$-2 \cdot x - \lambda_1 + \lambda_3 = 0$$
$$-2 \cdot y - \lambda_2 + \lambda_3 = 0$$
$$\lambda_1, \lambda_2, \lambda_3 \geq 0$$
$$\lambda_1 \cdot x = 0$$
$$\lambda_2 \cdot y = 0$$
$$\lambda_3 \cdot (1 - x - y) = 0$$

Stationary points

If $\lambda_3 = 0$, it can be shown that $x = y = 0$.

On the other hand, if $\lambda_3 \neq 0$, then $x + y = 1$, hence the problem is reduced to one dimensional case, which by standard methods lead to points $(x = 0, y = 1)$, $(x = 1, y = 0)$ and $(x = y = \frac{1}{2})$.

Second order necessary conditions

Second order necessary conditions

Theorem

Second order necessary conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Second order necessary conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem,

Second order necessary conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Second order necessary conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Then there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the previous theorem as well as the following condition:

Second order necessary conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Then there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the previous theorem as well as the following condition:

$$\bigtriangledown^2 f(\mathbf{x}^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot \bigtriangledown^2 g_i(\mathbf{x}^*)$$

Is positive semidefinite on the tangent subspace of active constraints at $\mathbf{x}^*$.

Second order necessary conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Then there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the previous theorem as well as the following condition:

$$\bigtriangledown^2 f(\mathbf{x}^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot \bigtriangledown^2 g_i(\mathbf{x}^*)$$

Is positive semidefinite on the tangent subspace of active constraints at $\mathbf{x}^*$.

Active Constraint

Second order necessary conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a local minimizer of the non-linear problem, and assume that $\mathbf{x}^*$ is a regular point for the constraints of the problem.

Then there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the previous theorem as well as the following condition:

$$\bigtriangledown^2 f(\mathbf{x}^*) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \cdot \bigtriangledown^2 g_i(\mathbf{x}^*)$$

Is positive semidefinite on the tangent subspace of active constraints at $\mathbf{x}^*$.

Active Constraint

Recall that a constraint is said to be active at a point $\mathbf{x}^*$, if it satisfies the constraint with equality.

Second order necessary conditions

Second order necessary conditions

Restatement of the last condition

Second order necessary conditions

Restatement of the last condition

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$,

Second order necessary conditions

Restatement of the last condition

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

Second order necessary conditions

Restatement of the last condition

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

Then, the last condition of the previous theorem, is equivalent to the following condition:

Second order necessary conditions

Restatement of the last condition

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

Then, the last condition of the previous theorem, is equivalent to the following condition:

$$\mathbf{N}^{\mathbf{T}}(\mathbf{x}^*) \cdot (\bigtriangledown^{\mathbf{2}}\mathbf{f}(\mathbf{x}^*) - \sum_{\mathbf{i} \in \mathcal{E} \cup \mathcal{I}} \lambda_{\mathbf{i}} \cdot \bigtriangledown^{\mathbf{2}}\mathbf{g}_{\mathbf{i}}(\mathbf{x}^*)) \cdot \mathbf{N}(\mathbf{x}^*)$$

Is positive semidefinite.

Second order sufficient conditions

Second order sufficient conditions

Theorem

Second order sufficient conditions

Theorem

Assume that f *and* g_i, $i \in \mathcal{E} \cup \mathcal{I}$ *are all twice continuously differentiable functions.*

Second order sufficient conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a feasible solution of the non-linear problem,

Second order sufficient conditions

Theorem

Assume that f *and* g_i, $i \in \mathcal{E} \cup \mathcal{I}$ *are all twice continuously differentiable functions.*

Let $\mathbf{x}^*$ *be a feasible solution of the non-linear problem, and assume that it is a regular point for the constraints of the problem.*

Second order sufficient conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a feasible solution of the non-linear problem, and assume that it is a regular point for the constraints of the problem.

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

Second order sufficient conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a feasible solution of the non-linear problem, and assume that it is a regular point for the constraints of the problem.

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

If there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the first order necessary theorem as well as the following condition:

Second order sufficient conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a feasible solution of the non-linear problem, and assume that it is a regular point for the constraints of the problem.

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

If there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the first order necessary theorem as well as the following condition:

$$g_i(\mathbf{x}^*) = 0, i \in \mathcal{I} \text{ implies } \lambda_i > 0,$$

Second order sufficient conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a feasible solution of the non-linear problem, and assume that it is a regular point for the constraints of the problem.

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

If there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the first order necessary theorem as well as the following condition:

$$g_i(\mathbf{x}^*) = 0, i \in \mathcal{I} \text{ implies } \lambda_i > 0,$$

and

$$\mathbf{N}^{\mathbf{T}}(\mathbf{x}^*) \cdot (\bigtriangledown^2 \mathbf{f}(\mathbf{x}^*) - \sum_{\mathbf{i} \in \mathcal{E} \cup \mathcal{I}} \lambda_{\mathbf{i}} \cdot \bigtriangledown^2 \mathbf{g}_{\mathbf{i}}(\mathbf{x}^*)) \cdot \mathbf{N}(\mathbf{x}^*)$$

is positive semidefinite,

Second order sufficient conditions

Theorem

Assume that f and g_i, $i \in \mathcal{E} \cup \mathcal{I}$ are all twice continuously differentiable functions.

Let $\mathbf{x}^*$ be a feasible solution of the non-linear problem, and assume that it is a regular point for the constraints of the problem.

Let $\mathbf{A}(\mathbf{x}^*)$ denote the Jacobian of the active constraints at $\mathbf{x}^*$, and let $\mathbf{N}(\mathbf{x}^*)$ be a null-space basis for $\mathbf{A}(\mathbf{x}^*)$.

If there exists $\lambda_i, i \in \mathcal{E} \cup \mathcal{I}$ satisfying the conditions of the first order necessary theorem as well as the following condition:

$$g_i(\mathbf{x}^*) = 0, i \in \mathcal{I} \textit{ implies } \lambda_i > 0,$$

and

$$\mathbf{N^T}(\mathbf{x}^*) \cdot (\bigtriangledown^2 \mathbf{f}(\mathbf{x}^*) - \sum_{\mathbf{i} \in \mathcal{E} \cup \mathcal{I}} \lambda_\mathbf{i} \cdot \bigtriangledown^2 \mathbf{g_i}(\mathbf{x}^*)) \cdot \mathbf{N}(\mathbf{x}^*)$$

is positive semidefinite, then $\mathbf{x}^*$ is local minimizer of the non-linear program.

Constrained non-linear programs

Constrained non-linear programs

An approach

Constrained non-linear programs

An approach

Below, we introduce an approach for solving non-linear programs.

Constrained non-linear programs

An approach

Below, we introduce an approach for solving non-linear programs.

It relies on the method of steepest decent method.

Constrained non-linear programs

An approach

Below, we introduce an approach for solving non-linear programs.

It relies on the method of steepest decent method.

The idea is to reduce the number of variables using the constraints and to solve this reduced and unconstrained problem using the steepest decent method.

Linear equality constraints

Linear equality constraints

Problem Formulation

Linear equality constraints

Problem Formulation

$$\min f(\mathbf{x}) := x_1^3 + x_2 + x_3^3 + x_4$$

Linear equality constraints

Problem Formulation

$$\min f(\mathbf{x}) := x_1^3 + x_2 + x_3^3 + x_4$$
$$g_1(\mathbf{x}) := x_1 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$

Linear equality constraints

Problem Formulation

$$\min f(\mathbf{x}) := x_1^3 + x_2 + x_3^3 + x_4$$
$$g_1(\mathbf{x}) := x_1 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$
$$g_2(\mathbf{x}) := -x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4 + 2 = 0,$$

Linear equality constraints

Linear equality constraints

Solving the linear equations

Linear equality constraints

Solving the linear equations

$$x_2 = 3 \cdot x_1 + 8 \cdot x_4 - 8,$$

Linear equality constraints

Solving the linear equations

$$x_2 = 3 \cdot x_1 + 8 \cdot x_4 - 8,$$
$$x_3 = -x_1 - 3 \cdot x_4 + 3,$$

Linear equality constraints

Linear equality constraints

Solving an unconstrained non-linear program

Linear equality constraints

Solving an unconstrained non-linear program

The above equations lead to the following unconstrained non-linear program:

Linear equality constraints

Solving an unconstrained non-linear program

The above equations lead to the following unconstrained non-linear program:

$$\min f(x_1, x_4) := x_1^3 + (3 \cdot x_1 + 8 \cdot x_4 - 8) + (-x_1 - 3 \cdot x_4 + 3)^3 + x_4.$$

Non-linear equality constraints

Non-linear equality constraints

Solving an unconstrained non-linear program

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$
$$g_1(\mathbf{x}) := x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$

$$g_1(\mathbf{x}) := x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$

$$g_2(\mathbf{x}) := -x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$
$$g_1(\mathbf{x}) := x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$
$$g_2(\mathbf{x}) := -x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

In this example, the constraints are non-linear.

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$

$$g_1(\mathbf{x}) := x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$

$$g_2(\mathbf{x}) := -x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

In this example, the constraints are non-linear.

Taylor series

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$
$$g_1(\mathbf{x}) := x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$
$$g_2(\mathbf{x}) := -x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

In this example, the constraints are non-linear.

Taylor series

Approximate the constraint functions by their Taylor series:

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\begin{aligned} \min f(\mathbf{x}) &:= x_1^2 + x_2 + x_3^2 + x_4 \\ g_1(\mathbf{x}) &:= x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0, \\ g_2(\mathbf{x}) &:= -x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0, \end{aligned}$$

In this example, the constraints are non-linear.

Taylor series

Approximate the constraint functions by their Taylor series:
$\mathbf{g}(\mathbf{x}) \approx \mathbf{g}(\bar{\mathbf{x}}) + \bigtriangledown \mathbf{g}(\bar{\mathbf{x}}) \cdot (\mathbf{x} - \bar{\mathbf{x}})^T$,

Non-linear equality constraints

Solving an unconstrained non-linear program

Consider the following non-linear program similar to the previous one:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$
$$g_1(\mathbf{x}) := x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$
$$g_2(\mathbf{x}) := -x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

In this example, the constraints are non-linear.

Taylor series

Approximate the constraint functions by their Taylor series:
$\mathbf{g}(\mathbf{x}) \approx \mathbf{g}(\bar{\mathbf{x}}) + \bigtriangledown \mathbf{g}(\bar{\mathbf{x}}) \cdot (\mathbf{x} - \bar{\mathbf{x}})^T$, where $\bar{\mathbf{x}}$ is the current point.

Non-linear equality constraints

Non-linear equality constraints

Our constraints become

Non-linear equality constraints

Our constraints become

$$g_1(\mathbf{x}) \approx 2 \cdot \bar{x}_1 \cdot x_1 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - (\bar{x}_1^2 + 4) = 0,$$

Non-linear equality constraints

Our constraints become

$$g_1(\mathbf{x}) \approx 2 \cdot \bar{x}_1 \cdot x_1 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - (\bar{x}_1^2 + 4) = 0,$$
$$g_2(\mathbf{x}) \approx -x_1 + x_2 + 2 \cdot x_3 - 4 \cdot \bar{x}_4 \cdot x_4 + (\bar{x}_4^2 + 2) = 0,$$

Non-linear equality constraints

Non-linear equality constraints

The general idea

Non-linear equality constraints

The general idea

The idea of Generalized Reduced Gradient Method (GRG) is to solve a sequence of sub-problems, each of which uses a linear approximation of the constraints.

Non-linear equality constraints

The general idea

The idea of Generalized Reduced Gradient Method (GRG) is to solve a sequence of sub-problems, each of which uses a linear approximation of the constraints.

In each iteration of the algorithm, the constraint linearization is recalculated at the point found from the previous iteration.

Non-linear equality constraints

The general idea

The idea of Generalized Reduced Gradient Method (GRG) is to solve a sequence of sub-problems, each of which uses a linear approximation of the constraints.

In each iteration of the algorithm, the constraint linearization is recalculated at the point found from the previous iteration.

Though the constraints are only approximated, the subproblems yield points that are progressively closer to the optimal point.

Non-linear equality constraints

Non-linear equality constraints

Our example: Starting point

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

This point satisfies our constraints.

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

This point satisfies our constraints.

It is quite possible to start with an infeasible point.

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

This point satisfies our constraints.

It is quite possible to start with an infeasible point.

Our example: the resulting program

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

This point satisfies our constraints.

It is quite possible to start with an infeasible point.

Our example: the resulting program

Using the approximation formulas derived earlier, we get:

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

This point satisfies our constraints.

It is quite possible to start with an infeasible point.

Our example: the resulting program

Using the approximation formulas derived earlier, we get:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

This point satisfies our constraints.

It is quite possible to start with an infeasible point.

Our example: the resulting program

Using the approximation formulas derived earlier, we get:

$$\min f(\mathbf{x}) := x_1^2 + x_2 + x_3^2 + x_4$$
$$g_1(\mathbf{x}) := x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0,$$

Non-linear equality constraints

Our example: Starting point

Let us start with $\mathbf{x}^0 = (0, -8, 3, 0)$.

This point satisfies our constraints.

It is quite possible to start with an infeasible point.

Our example: the resulting program

Using the approximation formulas derived earlier, we get:

$$\begin{aligned} \min f(\mathbf{x}) &:= x_1^2 + x_2 + x_3^2 + x_4 \\ g_1(\mathbf{x}) &:= x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4 = 0, \\ g_2(\mathbf{x}) &:= -x_1 + x_2 + 2 \cdot x_3 + 2 = 0, \end{aligned}$$

Non-linear equality constraints

Non-linear equality constraints

Solving the equality constraints

Non-linear equality constraints

Solving the equality constraints

Solving with respect to x_2 and x_3, we get:

Non-linear equality constraints

Solving the equality constraints

Solving with respect to x_2 and x_3, we get:

$$x_2 = 2 \cdot x_1 + 4 \cdot x_4 - 8,$$

Non-linear equality constraints

Solving the equality constraints

Solving with respect to x_2 and x_3, we get:

$$x_2 = 2 \cdot x_1 + 4 \cdot x_4 - 8,$$

$$x_3 = -\frac{1}{2} \cdot x_1 - 2 \cdot x_4 + 3,$$

Non-linear equality constraints

Non-linear equality constraints

The resulting unconstrained program

Non-linear equality constraints

The resulting unconstrained program

$$\min f(x_1, x_4) := x_1^2 + (2 \cdot x_1 + 4 \cdot x_4 - 8) + (-\frac{1}{2} \cdot x_1 - 2 \cdot x_4 + 3)^2 + x_4.$$

Non-linear equality constraints

The resulting unconstrained program

$$\min f(x_1, x_4) := x_1^2 + (2 \cdot x_1 + 4 \cdot x_4 - 8) + (-\tfrac{1}{2} \cdot x_1 - 2 \cdot x_4 + 3)^2 + x_4.$$

The next point

Non-linear equality constraints

The resulting unconstrained program

$$\min f(x_1, x_4) := x_1^2 + (2 \cdot x_1 + 4 \cdot x_4 - 8) + (-\frac{1}{2} \cdot x_1 - 2 \cdot x_4 + 3)^2 + x_4.$$

The next point

Solving this unconstrained minimization problem, we get $x_1 = -0.375$ and $x_4 = 0.96875$.

Non-linear equality constraints

The resulting unconstrained program

$$\min f(x_1, x_4) := x_1^2 + (2 \cdot x_1 + 4 \cdot x_4 - 8) + (-\tfrac{1}{2} \cdot x_1 - 2 \cdot x_4 + 3)^2 + x_4.$$

The next point

Solving this unconstrained minimization problem, we get $x_1 = -0.375$ and $x_4 = 0.96875$.

Substituting in equations for x_2 and x_3 gives $x_2 = -4.875$ and $x_3 = 1.25$.

Non-linear equality constraints

The resulting unconstrained program

$$\min f(x_1, x_4) := x_1^2 + (2 \cdot x_1 + 4 \cdot x_4 - 8) + (-\frac{1}{2} \cdot x_1 - 2 \cdot x_4 + 3)^2 + x_4.$$

The next point

Solving this unconstrained minimization problem, we get $x_1 = -0.375$ and $x_4 = 0.96875$.

Substituting in equations for x_2 and x_3 gives $x_2 = -4.875$ and $x_3 = 1.25$.

Thus, the iteration of GRG method is $\mathbf{x}^1 = (-0.375, -4.875, 1.25, 0.96875)$.

Non-linear equality constraints

Non-linear equality constraints

Continuing the process

Non-linear equality constraints

Continuing the process

To continue the solution process, we would re-linearize the constraint functions at the new point,

Non-linear equality constraints

Continuing the process

To continue the solution process, we would re-linearize the constraint functions at the new point,

Use the resulting system of linear equations to express two of the variables in terms of the others,

Non-linear equality constraints

Continuing the process

To continue the solution process, we would re-linearize the constraint functions at the new point,

Use the resulting system of linear equations to express two of the variables in terms of the others,

Substitute into the objective to get the new reduced problem,

Non-linear equality constraints

Continuing the process

To continue the solution process, we would re-linearize the constraint functions at the new point,

Use the resulting system of linear equations to express two of the variables in terms of the others,

Substitute into the objective to get the new reduced problem,

Solve the reduced problem for $\mathbf{x}^2$, and so forth.

Non-linear equality constraints

Continuing the process

To continue the solution process, we would re-linearize the constraint functions at the new point,

Use the resulting system of linear equations to express two of the variables in terms of the others,

Substitute into the objective to get the new reduced problem,

Solve the reduced problem for $\mathbf{x}^2$, and so forth.

Stopping criterion

Non-linear equality constraints

Continuing the process

To continue the solution process, we would re-linearize the constraint functions at the new point,

Use the resulting system of linear equations to express two of the variables in terms of the others,

Substitute into the objective to get the new reduced problem,

Solve the reduced problem for $\mathbf{x}^2$, and so forth.

Stopping criterion

The stopping criterion is $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| < T$, where T is a small constant.

Non-linear equality constraints

Non-linear equality constraints

Our example

Non-linear equality constraints

Our example

For example if we take, $T = 0.0025$ in the above example, we get $\mathbf{x}^k = (-0.498, -4, 823, 1.534, 0.610)$ and $f(\mathbf{x}^k) = -1.612$.

Non-linear equality constraints

Our example

For example if we take, $T = 0.0025$ in the above example, we get $\mathbf{x}^k = (-0.498, -4, 823, 1.534, 0.610)$ and $f(\mathbf{x}^k) = -1.612$.

The optimum solution is $\mathbf{x}^* = (-0.500, -4.825, 1.534, 0.610)$ and has an objective value of -1.612.

Non-linear equality constraints

Non-linear equality constraints

Remark

Non-linear equality constraints

Remark

During the iteration, the values of $f(\mathbf{x}^k)$ can sometimes be smaller than the minimum value.

Non-linear equality constraints

Remark

During the iteration, the values of $f(\mathbf{x}^k)$ can sometimes be smaller than the minimum value.

How this is possible?

Non-linear equality constraints

Remark

During the iteration, the values of $f(\mathbf{x}^k)$ can sometimes be smaller than the minimum value.

How this is possible?

The reason is that the points $\mathbf{x}^k$ computed by GRG are usually not feasible.

Non-linear equality constraints

Remark

During the iteration, the values of $f(\mathbf{x}^k)$ can sometimes be smaller than the minimum value.

How this is possible?

The reason is that the points $\mathbf{x}^k$ computed by GRG are usually not feasible.

They are only feasible for linear approximations of these constraints.

Non-linear equality constraints

Non-linear equality constraints

Starting from an infeasible solution

Non-linear equality constraints

Starting from an infeasible solution

Assume that we have chosen the point $\mathbf{x}^0$ so that it is infeasible.

Non-linear equality constraints

Starting from an infeasible solution

Assume that we have chosen the point $\mathbf{x}^0$ so that it is infeasible.

We consider a phase 1 problem, which is the construction of a feasible solution.

Non-linear equality constraints

Starting from an infeasible solution

Assume that we have chosen the point $\mathbf{x}^0$ so that it is infeasible.

We consider a phase 1 problem, which is the construction of a feasible solution.

The objective function for the phase 1 problem is the sum of the absolute values of the violated constraints.

Non-linear equality constraints

Starting from an infeasible solution

Assume that we have chosen the point $\mathbf{x}^0$ so that it is infeasible.

We consider a phase 1 problem, which is the construction of a feasible solution.

The objective function for the phase 1 problem is the sum of the absolute values of the violated constraints.

The constraints for the phase 1 problem are the non-violated ones.

Non-linear equality constraints

Non-linear equality constraints

Starting from an infeasible solution: Our example

Non-linear equality constraints

Starting from an infeasible solution: Our example

If we had started from the point $\mathbf{x}^0 = (1, 1, 0, 1)$, which happens to be infeasible, then the phase 1 problem would be the following:

Non-linear equality constraints

Starting from an infeasible solution: Our example

If we had started from the point $\mathbf{x}^0 = (1, 1, 0, 1)$, which happens to be infeasible, then the phase 1 problem would be the following:

$$\min |x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4|$$

Non-linear equality constraints

Starting from an infeasible solution: Our example

If we had started from the point $\mathbf{x}^0 = (1, 1, 0, 1)$, which happens to be infeasible, then the phase 1 problem would be the following:

$$\min |x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4|$$
$$-x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

Non-linear equality constraints

Starting from an infeasible solution: Our example

If we had started from the point $\mathbf{x}^0 = (1, 1, 0, 1)$, which happens to be infeasible, then the phase 1 problem would be the following:

$$\min |x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4|$$
$$-x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

This is because $\mathbf{x}^0$ violates the first constraint and satisfies the second one.

Non-linear equality constraints

Starting from an infeasible solution: Our example

If we had started from the point $\mathbf{x}^0 = (1, 1, 0, 1)$, which happens to be infeasible, then the phase 1 problem would be the following:

$$\min |x_1^2 + x_2 + 4 \cdot x_3 + 4 \cdot x_4 - 4|$$
$$-x_1 + x_2 + 2 \cdot x_3 - 2 \cdot x_4^2 + 2 = 0,$$

This is because $\mathbf{x}^0$ violates the first constraint and satisfies the second one.

Observe that the value of the objective function is 0, if and only if the corresponding point is a feasible solution.

Linear inequality constraints

Linear inequality constraints

The general strategy

Linear inequality constraints

The general strategy

We will discuss how GRG solves problems when there are inequality as well as equality constraints.

Linear inequality constraints

The general strategy

We will discuss how GRG solves problems when there are inequality as well as equality constraints.

At each iteration, only the tight inequality constraints enter into the system of linear linear equations used for eliminating variables (active inequality constraints).

Linear inequality constraints

The general strategy

We will discuss how GRG solves problems when there are inequality as well as equality constraints.

At each iteration, only the tight inequality constraints enter into the system of linear linear equations used for eliminating variables (active inequality constraints).

The process is complicated by the fact that active inequality constraints at the current point may need to be released in order to move to a better solution.

Linear inequality constraints

Linear inequality constraints

An example

Linear inequality constraints

An example

$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$

Linear inequality constraints

An example

$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$
$$x_1 - x_2 \geq 0,$$

Linear inequality constraints

An example

$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$
$$x_1 - x_2 \geq 0,$$
$$x_1 \geq 0,$$

Linear inequality constraints

An example

$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$
$$x_1 - x_2 \geq 0,$$
$$x_1 \geq 0,$$
$$x_2 \geq 0,$$

Linear inequality constraints

An example

$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$
$$x_1 - x_2 \geq 0,$$
$$x_1 \geq 0,$$
$$x_2 \geq 0,$$
$$x_2 \leq 2,$$

Linear inequality constraints

An example

$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$
$$x_1 - x_2 \geq 0,$$
$$x_1 \geq 0,$$
$$x_2 \geq 0,$$
$$x_2 \leq 2,$$

The process

Linear inequality constraints

An example

$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$
$$x_1 - x_2 \geq 0,$$
$$x_1 \geq 0,$$
$$x_2 \geq 0,$$
$$x_2 \leq 2,$$

The process

Assume that the initial feasible solution is $\mathbf{x}^0 = (1, 0)$.

Linear inequality constraints

An example

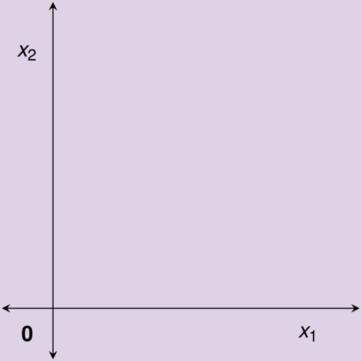
$$\min f(x_1, x_2) = (x_1 - \frac{1}{2})^2 + (x_2 - \frac{5}{2})^2$$
$$x_1 - x_2 \geq 0,$$
$$x_1 \geq 0,$$
$$x_2 \geq 0,$$
$$x_2 \leq 2,$$

The process

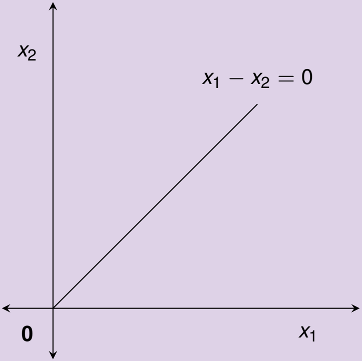
Assume that the initial feasible solution is $\mathbf{x}^0 = (1, 0)$.

It can be checked directly that the constraints $x_1 - x_2 \geq 0$, $x_1 \geq 0$, and $x_2 \leq 2$ are inactive, whereas the constraint $x_2 \geq 0$ is active.

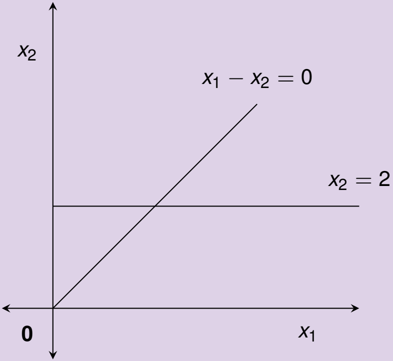
The feasible region



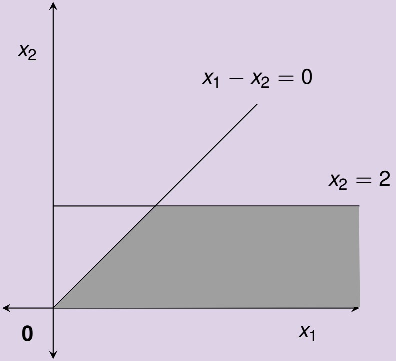
The feasible region



The feasible region



The feasible region



Linear inequality constraints

Linear inequality constraints

An example

Linear inequality constraints

An example

First let us evaluate the gradient of the objective function at $\mathbf{x}^0 = (1, 0)$.

Linear inequality constraints

An example

First let us evaluate the gradient of the objective function at $\mathbf{x}^0 = (1, 0)$.

$$\bigtriangledown f(\mathbf{x}^0) = (2 \cdot x_1^0 - 1, 2 \cdot x_2^0 - 5) = (1, -5).$$

Linear inequality constraints

An example

First let us evaluate the gradient of the objective function at $\mathbf{x}^0 = (1, 0)$.

$$\bigtriangledown f(\mathbf{x}^0) = (2 \cdot x_1^0 - 1, 2 \cdot x_2^0 - 5) = (1, -5).$$

This means that we will get the largest decrease in f if we move in the direction $\mathbf{d^0} = -\bigtriangledown \mathbf{f}(\mathbf{x^0}) = (-\mathbf{1}, \mathbf{5})$, that is, if we decrease x_1 and increase x_2.

Linear inequality constraints

An example

First let us evaluate the gradient of the objective function at $\mathbf{x}^0 = (1, 0)$.

$$\bigtriangledown f(\mathbf{x}^0) = (2 \cdot x_1^0 - 1, 2 \cdot x_2^0 - 5) = (1, -5).$$

This means that we will get the largest decrease in f if we move in the direction $\mathbf{d^0} = -\bigtriangledown \mathbf{f}(\mathbf{x^0}) = (-\mathbf{1}, \mathbf{5})$, that is, if we decrease x_1 and increase x_2.

The new point will be $\mathbf{x}^1 = \mathbf{x}^0 + \alpha^0 \cdot \mathbf{d^0}$ for some $\alpha^0 > 0$.

Linear inequality constraints

An example

First let us evaluate the gradient of the objective function at $\mathbf{x}^0 = (1, 0)$.

$$\bigtriangledown f(\mathbf{x}^0) = (2 \cdot x_1^0 - 1, 2 \cdot x_2^0 - 5) = (1, -5).$$

This means that we will get the largest decrease in f if we move in the direction $\mathbf{d^0} = -\bigtriangledown \mathbf{f}(\mathbf{x^0}) = (-\mathbf{1}, \mathbf{5})$, that is, if we decrease x_1 and increase x_2.

The new point will be $\mathbf{x}^1 = \mathbf{x}^0 + \alpha^0 \cdot \mathbf{d^0}$ for some $\alpha^0 > 0$.

The constraints imply that $\alpha^0 \leq 0.8333$.

Linear inequality constraints

An example

First let us evaluate the gradient of the objective function at $\mathbf{x}^0 = (1, 0)$.

$$\triangledown f(\mathbf{x}^0) = (2 \cdot x_1^0 - 1, 2 \cdot x_2^0 - 5) = (1, -5).$$

This means that we will get the largest decrease in f if we move in the direction $\mathbf{d^0} = -\triangledown \mathbf{f}(\mathbf{x^0}) = (-\mathbf{1}, \mathbf{5})$, that is, if we decrease x_1 and increase x_2.

The new point will be $\mathbf{x}^1 = \mathbf{x}^0 + \alpha^0 \cdot \mathbf{d^0}$ for some $\alpha^0 > 0$.

The constraints imply that $\alpha^0 \leq 0.8333$.

Now we perform a line search to determine the best value of α^0 in this range.

Linear inequality constraints

An example

First let us evaluate the gradient of the objective function at $\mathbf{x}^0 = (1, 0)$.

$$\bigtriangledown f(\mathbf{x}^0) = (2 \cdot x_1^0 - 1, 2 \cdot x_2^0 - 5) = (1, -5).$$

This means that we will get the largest decrease in f if we move in the direction $\mathbf{d^0} = -\bigtriangledown \mathbf{f}(\mathbf{x^0}) = (-\mathbf{1}, \mathbf{5})$, that is, if we decrease x_1 and increase x_2.

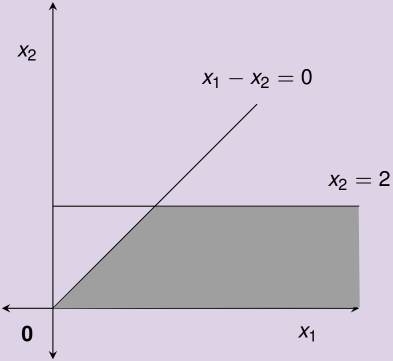
The new point will be $\mathbf{x}^1 = \mathbf{x}^0 + \alpha^0 \cdot \mathbf{d^0}$ for some $\alpha^0 > 0$.

The constraints imply that $\alpha^0 \leq 0.8333$.

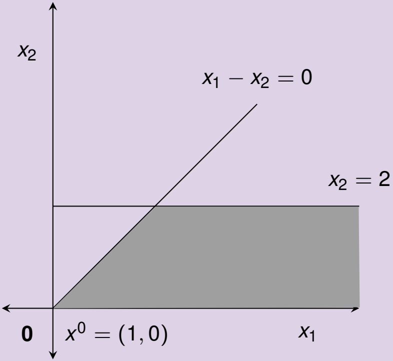
Now we perform a line search to determine the best value of α^0 in this range.

It can be shown that $\alpha^0 = 0.8333$, so $\mathbf{x}^1 = (0.8333, 0.8333)$.

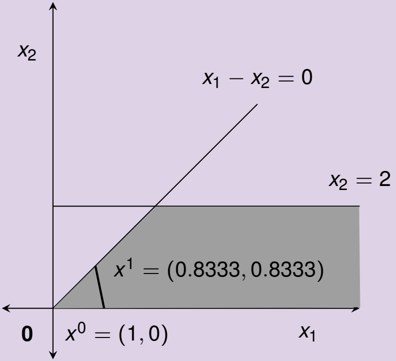
The feasible region and the progress of the algorithm



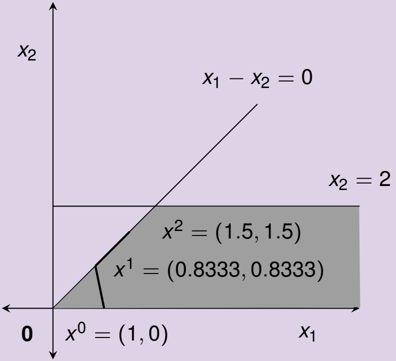
The feasible region and the progress of the algorithm



The feasible region and the progress of the algorithm



The feasible region and the progress of the algorithm



Non-linear programming

Non-linear programming

Problem Formulation

Non-linear programming

Problem Formulation

Consider a general non-linear optimization problem:

$$\min f(\mathbf{x})$$

Non-linear programming

Problem Formulation

Consider a general non-linear optimization problem:

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$

Non-linear programming

Problem Formulation

Consider a general non-linear optimization problem:

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Non-linear programming

Problem Formulation

Consider a general non-linear optimization problem:

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Quadratic Programs

Non-linear programming

Problem Formulation

Consider a general non-linear optimization problem:

$$\min f(\mathbf{x})$$
$$g_i(\mathbf{x}) = 0, i \in \mathcal{E}$$
$$g_i(\mathbf{x}) \geq 0, i \in \mathcal{I}$$

Quadratic Programs

In order to solve this problem, we can use methods available for quadratic programs.

Non-linear programming

Problem Formulation

Consider a general non-linear optimization problem:

$$\begin{aligned} \min f(\mathbf{x}) \\ g_i(\mathbf{x}) = 0, i \in \mathcal{E} \\ g_i(\mathbf{x}) \geq 0, i \in \mathcal{I} \end{aligned}$$

Quadratic Programs

In order to solve this problem, we can use methods available for quadratic programs.

This is the idea behind sequential quadratic programming.

Non-linear programming

Non-linear programming

The idea

Non-linear programming

The idea

At the current feasible point $\mathbf{x}^k$, the problem is approximated by a quadratic program:

Non-linear programming

The idea

At the current feasible point $\mathbf{x}^k$, the problem is approximated by a quadratic program:

A quadratic approximation of the Lagrangian function is computed as well as linear approximations of the constraints.

Non-linear programming

The idea

At the current feasible point $\mathbf{x}^k$, the problem is approximated by a quadratic program:

A quadratic approximation of the Lagrangian function is computed as well as linear approximations of the constraints.

The resulting quadratic program is of the form:

$$\min(\triangledown f(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{x}^k)^T \cdot \mathbf{B_k} \cdot (\mathbf{x} - \mathbf{x^k}))$$

Non-linear programming

The idea

At the current feasible point $\mathbf{x}^k$, the problem is approximated by a quadratic program:

A quadratic approximation of the Lagrangian function is computed as well as linear approximations of the constraints.

The resulting quadratic program is of the form:

$$\min(\triangledown f(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{x}^k)^T \cdot \mathbf{B_k} \cdot (\mathbf{x} - \mathbf{x^k}))$$

$$\triangledown g_i(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + g_i(\mathbf{x}^k) = 0, i \in \mathcal{E}$$

Non-linear programming

The idea

At the current feasible point $\mathbf{x}^k$, the problem is approximated by a quadratic program:

A quadratic approximation of the Lagrangian function is computed as well as linear approximations of the constraints.

The resulting quadratic program is of the form:

$$\min(\bigtriangledown f(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{x}^k)^T \cdot \mathbf{B_k} \cdot (\mathbf{x} - \mathbf{x^k}))$$

$$\bigtriangledown g_i(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + g_i(\mathbf{x}^k) = 0, i \in \mathcal{E}$$

$$\bigtriangledown g_i(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + g_i(\mathbf{x}^k) \geq 0, i \in \mathcal{I}$$

Non-linear programming

The idea

At the current feasible point $\mathbf{x}^k$, the problem is approximated by a quadratic program:

A quadratic approximation of the Lagrangian function is computed as well as linear approximations of the constraints.

The resulting quadratic program is of the form:

$$\min(\bigtriangledown f(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + \frac{1}{2} \cdot (\mathbf{x} - \mathbf{x}^k)^T \cdot \mathbf{B_k} \cdot (\mathbf{x} - \mathbf{x^k}))$$

$$\bigtriangledown g_i(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + g_i(\mathbf{x}^k) = 0, i \in \mathcal{E}$$

$$\bigtriangledown g_i(\mathbf{x}^k)^T \cdot (\mathbf{x} - \mathbf{x}^k) + g_i(\mathbf{x}^k) \geq 0, i \in \mathcal{I}$$

Where $\mathbf{B_k} = \bigtriangledown^2_{\mathbf{xx}} \mathbf{L}(\mathbf{x^k}, \lambda^\mathbf{k})$ is the Hessian of the Langrangian function with respect to the variables x and λ^k is the current estimate of the Lagrangian multipliers.

Non-linear programming

Non-linear programming

Next iteration

Non-linear programming

Next iteration

This problem can be solved with one of the methods developed for quadratic programs.

Non-linear programming

Next iteration

This problem can be solved with one of the methods developed for quadratic programs.

The optimal solution of the quadratic program is used to determine a search direction.

Non-linear programming

Next iteration

This problem can be solved with one of the methods developed for quadratic programs.

The optimal solution of the quadratic program is used to determine a search direction.

Then a line search is performed to determine the next iterate.

Non-smooth Optimization

Non-smooth Optimization

Problem Formulation

Non-smooth Optimization

Problem Formulation

We will consider unconstrained nonlinear programs of the form

$$\min f(\mathbf{x})$$

Non-smooth Optimization

Problem Formulation

We will consider unconstrained nonlinear programs of the form

$$\min f(\mathbf{x})$$

where $\mathbf{x} = (x_1, ..., x_n)$ and f is a non-differentiable convex function.

Non-smooth Optimization

Problem Formulation

We will consider unconstrained nonlinear programs of the form

$$\min f(\mathbf{x})$$

where $\mathbf{x} = (x_1, ..., x_n)$ and f is a non-differentiable convex function.

Optimality conditions based on the gradient are not available since the gradient is not always defined.

Non-smooth Optimization

Problem Formulation

We will consider unconstrained nonlinear programs of the form

$$\min f(\mathbf{x})$$

where $\mathbf{x} = (x_1, ..., x_n)$ and f is a non-differentiable convex function.

Optimality conditions based on the gradient are not available since the gradient is not always defined.

However, the notion of gradient can be generalized as follows.

Non-smooth Optimization

Problem Formulation

We will consider unconstrained nonlinear programs of the form

$$\min f(\mathbf{x})$$

where $\mathbf{x} = (x_1, ..., x_n)$ and f is a non-differentiable convex function.

Optimality conditions based on the gradient are not available since the gradient is not always defined.

However, the notion of gradient can be generalized as follows.

Definition

Non-smooth Optimization

Problem Formulation

We will consider unconstrained nonlinear programs of the form

$$\min f(\mathbf{x})$$

where $\mathbf{x} = (x_1, ..., x_n)$ and f is a non-differentiable convex function.

Optimality conditions based on the gradient are not available since the gradient is not always defined.

However, the notion of gradient can be generalized as follows.

Definition

A subgradient of f at point $\mathbf{x}^*$ is a vector $\mathbf{s}^* = (\mathbf{s}_1^*, ..., \mathbf{s}_n^*)$ such that

Non-smooth Optimization

Problem Formulation

We will consider unconstrained nonlinear programs of the form

$$\min f(\mathbf{x})$$

where $\mathbf{x} = (x_1, ..., x_n)$ and f is a non-differentiable convex function.

Optimality conditions based on the gradient are not available since the gradient is not always defined.

However, the notion of gradient can be generalized as follows.

Definition

A subgradient of f at point $\mathbf{x}^*$ is a vector $\mathbf{s}^* = (\mathbf{s}_1^*, ..., \mathbf{s}_n^*)$ such that

$$\mathbf{s}^* \cdot (\mathbf{x} - \mathbf{x}^*) \leq f(\mathbf{x}) - f(\mathbf{x}^*), \text{ for every } \mathbf{x}.$$

Non-smooth Optimization

Non-smooth Optimization

Subgradients

Non-smooth Optimization

Subgradients

When the function is differentiable, the subgradient is identical to the gradient.

Non-smooth Optimization

Subgradients

When the function is differentiable, the subgradient is identical to the gradient.

When f is not differentiable at point $\mathbf{x}$, there are typically many subgradients at $\mathbf{x}$.

Non-smooth Optimization

Subgradients

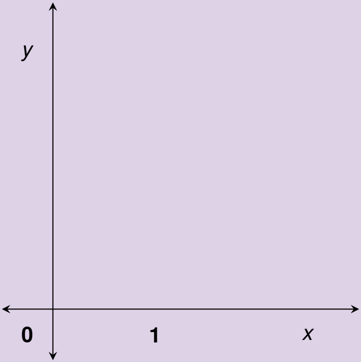
When the function is differentiable, the subgradient is identical to the gradient.

When f is not differentiable at point $\mathbf{x}$, there are typically many subgradients at $\mathbf{x}$.

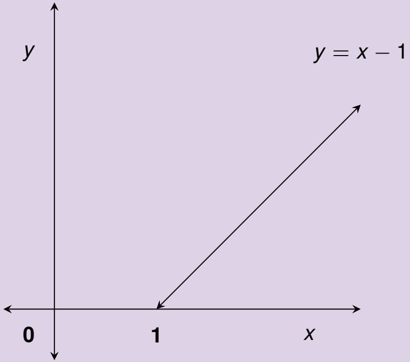
For example consider the function

$$f(x) = |x - 1|.$$

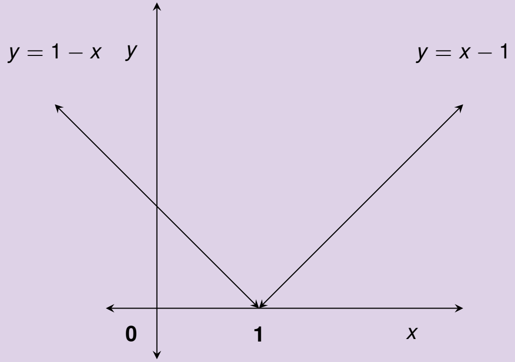
The subgradients of the function $y = |x - 1|$



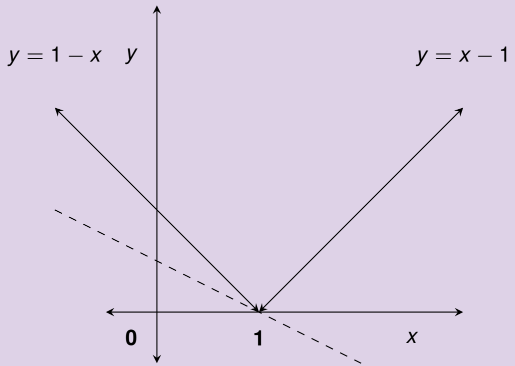
The subgradients of the function $y = |x - 1|$



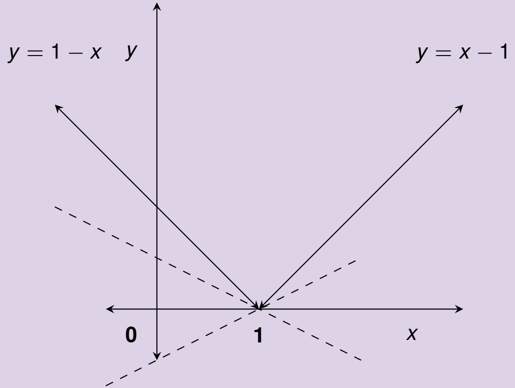
The subgradients of the function $y = |x - 1|$



The subgradients of the function $y = |x - 1|$



The subgradients of the function $y = |x - 1|$



A property of convex functions

A property of convex functions

Theorem

A property of convex functions

Theorem

Let f is a non-differentiable convex function.

A property of convex functions

Theorem

Let f is a non-differentiable convex function.

The point $\mathbf{x}^*$ is a minimum of f, if and only if f has a zero subgradient at $\mathbf{x}^*$.

A property of convex functions

Theorem

Let f is a non-differentiable convex function.

The point $\mathbf{x}^*$ is a minimum of f, if and only if f has a zero subgradient at $\mathbf{x}^*$.

Example

A property of convex functions

Theorem

Let f is a non-differentiable convex function.

The point $\mathbf{x}^*$ is a minimum of f, if and only if f has a zero subgradient at $\mathbf{x}^*$.

Example

In the case of $f(x) = |x - 1|$, 0 is a subgradient at point $x^* = 1$, therefore this is the point where the minimum of f is achieved.

The method of steepest decent for convex functions

The method of steepest decent for convex functions

Idea

The method of steepest decent for convex functions

Idea

The method of steepest decent can be extended to non-differentiable functions.

The method of steepest decent for convex functions

Idea

The method of steepest decent can be extended to non-differentiable functions.

First we compute any subgradient direction at the current point, and use its opposite direction to make the next step.

The method of steepest decent for convex functions

Idea

The method of steepest decent can be extended to non-differentiable functions.

First we compute any subgradient direction at the current point, and use its opposite direction to make the next step.

Though subgradient directions are not always directions of ascent, one can still guarantee convergence to the optimum point by choosing the step size appropriately.

The method of steepest decent for convex functions

The method of steepest decent for convex functions

Formal description of the method

1: Input: Any point $\mathbf{x}^0$.

The method of steepest decent for convex functions

Formal description of the method

1: Input: Any point $\mathbf{x}^0$.
2: Output: A point $\mathbf{x}^k$.

The method of steepest decent for convex functions

Formal description of the method

1: Input: Any point $\mathbf{x}^0$.
2: Output: A point $\mathbf{x}^k$.
3: Set $i = 0$.

The method of steepest decent for convex functions

Formal description of the method

1: Input: Any point $\mathbf{x}^0$.
2: Output: A point $\mathbf{x}^k$.
3: Set $i = 0$.
4: Compute a subgradient $\mathbf{s^i}$ of f at point $\mathbf{x}^i$.

The method of steepest decent for convex functions

Formal description of the method

1: Input: Any point $\mathbf{x}^0$.
2: Output: A point $\mathbf{x}^k$.
3: Set $i = 0$.
4: Compute a subgradient $\mathbf{s}^{\mathbf{i}}$ of f at point $\mathbf{x}^i$.
5: **if** ($\mathbf{s}^{\mathbf{i}}$ is $\mathbf{0}$ or close to it) **then**

The method of steepest decent for convex functions

Formal description of the method

1: Input: Any point $\mathbf{x}^0$.
2: Output: A point $\mathbf{x}^k$.
3: Set $i = 0$.
4: Compute a subgradient $\mathbf{s^i}$ of f at point $\mathbf{x}^i$.
5: **if** ($\mathbf{s^i}$ is $\mathbf{0}$ or close to it) **then**
6: stop.

The method of steepest decent for convex functions

Formal description of the method

```
Input: Any point x^0.
Output: A point x^k.
Set i = 0.
Compute a subgradient s^i of f at point x^i.
if (s^i is 0 or close to it) then
    stop.
else
```

The method of steepest decent for convex functions

Formal description of the method

Input: Any point $\mathbf{x}^0$.
Output: A point $\mathbf{x}^k$.
Set $i = 0$.
Compute a subgradient $\mathbf{s^i}$ of f at point $\mathbf{x}^i$.
if ($\mathbf{s^i}$ is $\mathbf{0}$ or close to it) **then**
    stop.
else
    Let $\mathbf{x^{i+1}} = \mathbf{x}^i - \alpha_i \cdot \mathbf{s^i}$, where $\alpha_i > 0$ denotes the step size.

The method of steepest decent for convex functions

Formal description of the method

Input: Any point $\mathbf{x}^0$.
Output: A point $\mathbf{x}^k$.
Set $i = 0$.
Compute a subgradient $\mathbf{s}^{\mathbf{i}}$ of f at point $\mathbf{x}^i$.
if ($\mathbf{s}^{\mathbf{i}}$ is **0** or close to it) **then**
    stop.
else
    Let $\mathbf{x}^{\mathbf{i+1}} = \mathbf{x}^i - \alpha_i \cdot \mathbf{s}^{\mathbf{i}}$, where $\alpha_i > 0$ denotes the step size.
    Perform the next iteration.

The method of steepest decent for convex functions

Formal description of the method

Input: Any point $\mathbf{x}^0$.
Output: A point $\mathbf{x}^k$.
Set $i = 0$.
Compute a subgradient $\mathbf{s^i}$ of f at point $\mathbf{x}^i$.
if ($\mathbf{s^i}$ is $\mathbf{0}$ or close to it) **then**
    stop.
else
    Let $\mathbf{x^{i+1}} = \mathbf{x}^i - \alpha_i \cdot \mathbf{s^i}$, where $\alpha_i > 0$ denotes the step size.
    Perform the next iteration.
end if

The method of steepest decent for convex functions

The method of steepest decent for convex functions

Choice of α_i

The method of steepest decent for convex functions

Choice of α_i

Several choices of step size α_i have been proposed in the literature.

The method of steepest decent for convex functions

Choice of α_i

Several choices of step size α_i have been proposed in the literature.

To guarantee convergence to the optimum, the step size α_i needs to be decreased very slowly.

The method of steepest decent for convex functions

Choice of α_i

Several choices of step size α_i have been proposed in the literature.

To guarantee convergence to the optimum, the step size α_i needs to be decreased very slowly.

For example, for the choice of $\alpha_i \to 0$ such that $\sum_i \alpha_i = +\infty$, will do.

Literature

Literature

The list of references

Literature

The list of references

1. G. Cornuejols, R. Tütüncü, Optimization methods in Finance, Cambridge University Press, 2007.

Literature

The list of references

1. G. Cornuejols, R. Tütüncü, Optimization methods in Finance, Cambridge University Press, 2007.
2. M. Bartholomew-Biggs, Nonlinear Optimization with Financial Applications, Kluwer Academic Publishers, 2005.