Nonlinear Optimization

Linear Programming — The Simplex Method

Niclas Börlin

Department of Computing Science

Umeå University
niclas.borlin@cs.umu.se

December 10, 2007

Linear programming

"Interest in linear programming (LP) in its own right began in the late 1940's with the invention of the simplex method, and has continued unabated until the present day. Linear programming is viewed and taught in an astonishing variety of ways, to the extent that even an expert may be puzzled by someone else's version of the subject!"

Gill, Murray, Wright, "Numerical Linear Algebra and Optimization, vol. 1", Addison-Wesley, 1991, Chap. 7.

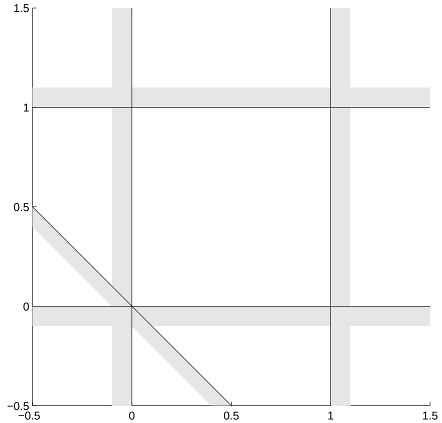
Problem formulation

- An *LP problem* (Linear Programming problem) is an optimization problem where the objective function and the constraints are linear functions.
- An example of an LP problem is

$$
\begin{array}{ccccc}
 & & \displaystyle\max_{x_1,x_2} & x_2 \\
 & & \text{s.t.} & \\
x_1 & & \geq & 0 \\
 & x_2 & \geq & 0 \\
x_1 & +x_2 & \geq & 0 \\
x_1 & & \leq & 1 \\
 & x_2 & \leq & 1
\end{array}
$$

Feasible region

The feasible region for the problem is



Constraints

- We have two different types of constraints; equality and inequality.
- A linear inequality constraint may be written as
$$a^T x \geq \beta$$
for some vector a.
- A point $\bar{x}$ is feasible with respect to the constraint $a^T x \geq \beta$ if the constrain is satisfied, i.e. $a^T \bar{x} \geq \beta$.
- If the constraint is satisfied with equality, the constraint is *active*, otherwise *inactive*.
- We will only work with $\geq$ inequality constraints, since
  - A constraint $a^T x \leq \beta$ may be rewritten as $-a^T x \geq -\beta$.
  - A constraint $a^T x = \beta$ may be described as two inequality constraints $a^T x \geq \beta$ and $a^T x \leq \beta$.

Standard form

- We will rewrite all LP problems to the following standard form:
$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \end{aligned}$$
- Maximization problems
$$\max_x c^T x$$
are rewritten as
$$\min_x -c^T x.$$

Our example problem

$$\begin{array}{cccc} & & \max_{x_1,x_2} & x_2 \\ & & \text{s.t.} & \\ x_1 & & \geq & 0 \\ & x_2 & \geq & 0 \\ x_1 & +x_2 & \geq & 0 \\ x_1 & & \leq & 1 \\ & x_2 & \leq & 1 \end{array}$$

becomes

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \end{aligned}$$

with

$$c = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \; A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \; b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ -1 \end{bmatrix}.$$

Sets of active/inactive constraints

- For a constraint $a_i^T x \geq \beta_i$ define the *residual* in the point $\bar{x}$ as
$$r_i(\bar{x}) = a_i^T \bar{x} - \beta_i.$$
- The residual is positive if the constraint is inactive, zero when the constraint is active, and negative when the constraint is violated.
- For the set of constraints $Ax \geq b$, the residual vector r is defined as
$$r(\bar{x}) = A\bar{x} - b.$$
- Given the inequality $Ax \geq b$ we define an index set $\mathcal{A}(\bar{x})$ as the set of indices
$$\{i : r_i(\bar{x}) = 0\},$$
i.e. the constraints that are satisfied with equality in the point $\bar{x}$.
- The matrix with the rows $\mathcal{A}(\bar{x})$ of A is denoted $A_{\mathcal{A}}(\bar{x})$ and called the *active set matrix*.

Algorithm outline

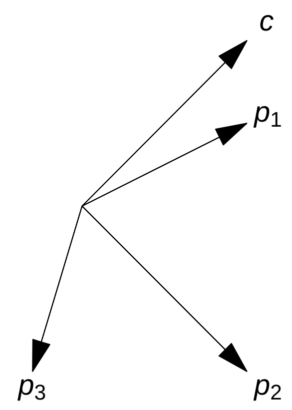
- We will adopt a linesearch-based descent algorithm for solving the LP.
  - Determine a starting approximation x_0.
  - Repeat for $k = 0, 1, \ldots$
    - If x_k optimal, terminate.
    - Determine a *search direction* p_k.
    - Determine a *step length* α_k.
    - $x_{k+1} = x_k + \alpha_k p_k$

Gradients and Descent directions

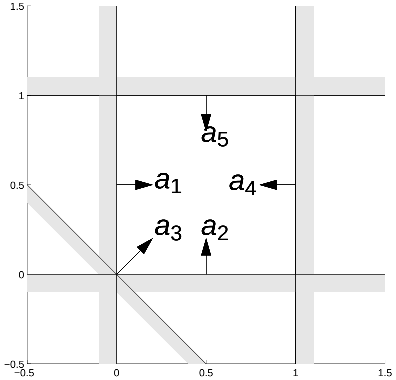
- The objective function of the minimization problem is written as
$$f(x) = c^T x,$$
where c is a constant vector.
- The gradient to $f(x)$ is
$$\nabla f(x) = c.$$
- Define $F(\alpha)$ as the function value starting from a point $\bar{x}$ along a search direction p and with a step length α:
$$F(\alpha) = f(\bar{x} + \alpha p) = c^T(\bar{x} + \alpha p) = c^T\bar{x} + \alpha c^T p = f(\bar{x}) + \alpha c^T p.$$
- Thus, the change in objective function value depends only on $c^T p$.

- $c^T p = 0$ means that
$$F(\alpha) = f(\bar{x}) + \alpha \underbrace{c^T p}_{=0} = f(\bar{x}),$$
i.e. the function value is constant.
- $c^T p < 0$ means that
$$F(\alpha) = f(\bar{x}) + \alpha \underbrace{c^T p}_{<0} < f(\bar{x}) \text{ for } \alpha > 0,$$
i.e. the function value is reduced.
- A search direction p such that $c^T p < 0$ is thus a *descent direction*.

In the figure below, $p_1^T c > 0$, $p_2^T c = 0$, and $p_3^T c < 0$.



- Each constraint splits $\Re^n$ in two halves of feasible points $\{x : a^T x \geq \beta\}$ and infeasible points $\{x : a^T x < \beta\}$.
- The gradient of a constraint $a^T x \geq \beta$ is a. The gradient is always directed *into* the feasible set.



- Consider a point $\bar{x}$ (not necessarily feasible), a direction p and a constraint $a_i^T x \geq \beta_i$.
- For a step α along p,

$$a_i^T(\bar{x} + \alpha p) = a_i^T \bar{x} + \alpha a_i^T p.$$

- The corresponding residual is

$$r_i(\bar{x} + \alpha p) = a_i^T \bar{x} + \alpha a_i^T p - \beta_i = r_i(\bar{x}) + \alpha a_i^T p.$$

- The corresponding residual is

$$r_i(\bar{x} + \alpha p) = a_i^T \bar{x} + \alpha a_i^T p - \beta_i = r_i(\bar{x}) + \alpha a_i^T p.$$

- Thus, if $a_i^T p = 0$, the residual value will not change, but if $a_i^T p \neq 0$ it is possible to calculate how long step σ_i we can take to satisfy and activate the constraint:

$$\begin{aligned} r_i(\bar{x} + \sigma_i p) &= 0 \\ r_i(\bar{x}) + \sigma_i a_i^T p &= 0 \end{aligned} \Rightarrow \sigma_i = \frac{r_i(\bar{x})}{-a_i^T p}$$

which is defined if $a_i^T p \neq 0$.
- For $a_i^T p = 0$,

$$\sigma_i = \begin{cases} +\infty & \text{if } r_i(\bar{x}) \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Search directions

- The vector $p \neq 0$ is a *feasible direction* in the feasible point $\bar{x}$ with respect to the constraint $a_i^T x \geq \beta_i$ if there exists a positive scalar τ_i such that
$$r_i(\bar{x} + \alpha p) \geq 0\ ,\ 0 \leq \alpha \leq \tau_i.$$
- For an inactive constraint ($r_i(\bar{x}) > 0$), all directions p are feasible, but for an active constraint ($r_i(\bar{x}) = 0$),

$$\begin{aligned} r_i(\bar{x} + \alpha p) &\geq 0, \\ a_i^T(\bar{x} + \alpha p) - \beta_i &= r_i(\bar{x}) + \alpha a_i^T p = \alpha a_i^T p \geq 0 \text{ for } \alpha > 0, \\ a_i^T p &\geq 0 \end{aligned}$$

- Thus, if $a_i^T p = 0$, the constraint remains active, if $a_i^T p > 0$ the constraint will be deactivated, and if $a_i^T p < 0$, the constraint will be violated.

- For a set of constraint, a vector p is a feasible direction from the feasible point $\bar{x}$ if $p \neq 0$ and there exists a positive scalar τ such that

$$\begin{aligned} r(\bar{x}+\alpha p) &\geq 0\ ,\ 0 \leq \alpha \leq \tau \\ &\Updownarrow \\ A(\bar{x}+\alpha p) &\geq b\ ,\ 0 \leq \alpha \leq \tau \end{aligned}$$

- As the only limitations on p are given by the active constraints, p will be feasible if

$$\begin{aligned} a_i^T p &\geq 0\ \ \forall i \in \mathcal{A}(\bar{x}) \\ &\Updownarrow \\ A_{\mathcal{A}} p &\geq 0, \end{aligned}$$

where $A_{\mathcal{A}}$ is the active set matrix in $\bar{x}$.

- If we solve

$$A_{\mathcal{A}}^T \lambda = c$$

and some λ_i is negative, a search direction p can be determined as the solution of

$$A_{\mathcal{A}} p = e_i,$$

where e_i is a zero vector except a 1 in the i:th element.
- This means that

$$r_i(\bar{x}+\alpha p) > 0 \text{ for } \alpha > 0,$$

i.e. constraint i will be deactivated.
- A p chosen this way will be a descent direction, since

$$c^T p = (A_{\mathcal{A}}^T \lambda)^T p = \lambda^T A_{\mathcal{A}} p = \lambda^T e_i = \lambda_i < 0.$$

Corners and edges

- A *corner* is an extreme point of the feasible set which cannot lie on a line between two other feasible points.
- A corner is restricted in all n dimensions, i.e. the active set matrix $A_{\mathcal{A}}$ has at least n linearly independent rows.
- If $A_{\mathcal{A}}$ has more than n rows, the corner is *degenerated*, otherwise *non-degenerated*.
- An *edge* is a linear subspace of $\Re^n$, and is defined by an active set matrix $A_{\mathcal{A}}$ with $n-1$ linearly independent rows.
- A corner is always a minimizer to an LP problem. If the minimizer is not unique, all points on an adjoining edge are minimizers.

Step length

- Assume that $\bar{x}$ is a corner of the feasible set and we have calculated a search direction p from $\bar{x}$.
- Let $\mathcal{D}$ denote the set of all inactive constraints with a diminishing residual along p, i.e. $j \in \mathcal{D}$ if $a_j^T p < 0$.
- Determine σ_j as the longest step that can be taken along p without violating constraint j, i.e.

$$\sigma_j = \begin{cases} \frac{-r_j(\bar{x})}{a_j^T p} & \text{if } j \in \mathcal{D} \\ +\infty & \text{otherwise} \end{cases}.$$

- Then

$$\alpha = \min_j \sigma_j$$

is the longest step we can take before activating *any* constraint.
- The new point is $\bar{x} + \alpha p$.

The Simplex Algorithm

Let $x^{[n]}$ denote the current point after n iterations. Given a non-degenerated corner $x^{[n]}$, the simplex algorithm is as follows:

- Determine a feasible descent direction p in $x^{[n]}$:
  - Identify the active set $A_{\mathcal{A}}$ in $\bar{x}$.
  - Solve
    $$A_{\mathcal{A}}^T \lambda = c.$$
  - Choose an i such that $\lambda_i < 0$.
  - Determine the search direction p as the solution of
    $$A_{\mathcal{A}} p = e_i.$$
- Determine the step length
  $$\alpha = \min_j \sigma_j,$$
  where
  $$\sigma_j = \begin{cases} \frac{-r_j(\bar{x})}{a_j^T p} & \text{if } j \in \mathcal{D} \\ +\infty & \text{otherwise} \end{cases}.$$
- $x^{[n+1]} = x^{[n]} + \alpha p$.