

The Affine Scaling Method

Overview

Given a linear programming problem in equality form with full rank constraint matrix and a strictly positive feasible solution x^0 , we transform the problem to a new one with feasible solution $\mathbf{1} = [1, 1, \dots, 1]^T$. The original problem and scaled problem

Original:	Scaled:
$\mathbf{max} \quad c^T x$	$\mathbf{max} \quad \hat{c}^T x$
$\mathbf{s. t.} \quad Ax = b$	$\mathbf{s. t.} \quad \hat{A}x = b$
$x \geq 0$	$x \geq 0$

are related as follows: X^0 is the diagonal matrix with the entries of x^0 down the diagonal; $\hat{A} = AX^0$; and $\hat{c} = X^0c$.

In the scaled problem, we

- project the gradient \hat{c} onto the subspace $\hat{A}x = 0$ to obtain a search direction d ;
- find the largest scalar β such that $\mathbf{1} + \beta d$ is still non-negative (the other feasibility conditions are already guaranteed by choice of d);
- obtain new (scaled) solution $\hat{x} = \mathbf{1} + r\beta d$ where $0 < r < 1$ is a global constant;
- update our solution x^0 to the next solution $x^1 = X^0\hat{x}$.

This process is repeated, as needed, until we are sufficiently close to an optimal solution (e.g., as one may check by solving for a dual vector using complementary slackness conditions).

One iteration of the affine scaling method

We assume we have an $m \times n$ matrix A with full row rank, an m -vector b and an n -vector c . (If A does not have full row rank, first row reduce, and if $Ax = b$ is feasible, then replace A by a smaller matrix.) We also have a fixed step size r , $0 < r < 1$.

Given a strictly positive feasible solution x^0 to the problem (so $Ax^0 = b$), we obtain a better strictly positive feasible solution x^1 as follows:

Step 1: (Scale) Let $\hat{A} = AX^0$ and $\hat{c} = X^0c$ where X^0 is the diagonal matrix with entries from x^0 down the diagonal. (Now we have a scaled problem with $\mathbf{1}$ as a feasible solution.)

Step 2: (Find projection) Compute the projection matrix onto the nullspace of \hat{A} :

$$P = I - \hat{A}^T (\hat{A}\hat{A}^T)^{-1} \hat{A}$$

Step 3: (Find search direction) The search direction (in the scaled problem) is then

$$d = P\hat{c}.$$

Step 4: (Compute ratios) Let $\beta = \min \left\{ -\frac{1}{d_j} : d_j < 0 \right\}$.

Step 5: (Scaled update) The improved solution in the scaled problem is

$$\hat{x} = \mathbf{1} + r\beta d$$

where r is the prescribed step size ($0 < r < 1$).

Step 6: (Next iterate) Now scale back to the original problem and take

$$x^1 = X^0\hat{x}.$$

Some explanations

In Step 1, we simply write down the data (\hat{A}, \hat{c}) for the scaled problem. Note that the right-hand side vector does not change. But that's okay because

$$\hat{A}\mathbf{1} = (AX^0)\mathbf{1} = A(X^0\mathbf{1}) = Ax^0 = b$$

so $\mathbf{1}$ is indeed a feasible solution to the scaled problem and it is “far away” from the boundaries $x_j = 0$. Moreover, X^0 is a diagonal matrix, so it is symmetric: $(X^0)^T = X^0$. Thus

$$\hat{c}^T\mathbf{1} = (X^0c)^T\mathbf{1} = (c^T(X^0)^T)\mathbf{1} = c^T(X^0\mathbf{1}) = c^T x^0.$$

More generally, the objective value of any solution \hat{x} to the scaled problem is equal to the objective value of the corresponding solution $x = X^0\hat{x}$ in the original problem.

In Step 2, we do the most work. We need to project the gradient \hat{c} onto the subspace $\{x : \hat{A}x = 0\}$ in order to have a feasible search direction. As we worked out in class, the projection matrix P does just the trick.

In Step 3, we take $d = P\hat{c}$ so that $\hat{A}d = 0$ and, for any scalar β , the vector $\hat{x} = \mathbf{1} + \beta d$ satisfies

$$\hat{A}\hat{x} = \hat{A}(\mathbf{1} + \beta d) = \hat{A}\mathbf{1} + \beta\hat{A}d = \hat{A}\mathbf{1} = b$$

as established above. So, no matter how far we travel in this direction, the equality constraints will still be satisfied.

So, in Step 4, we know that any vector of the form $\hat{x} = \mathbf{1} + \beta d$ satisfies the equality constraints. So in order to stay feasible, we simply need $\mathbf{1} + \beta d \geq 0$. This requires $1 + \beta d_j \geq 0$ for all j . Since we are going to take β positive (to improve the objective value), the only coordinates j that concern us are those with $d_j < 0$, and each of these gives us an upper limit $\beta \leq -1/d_j$.

In Step 5, we finally use our global step size r . We know that $\mathbf{1} + \beta d \geq 0$ and is otherwise feasible. But we need a *strictly positive* feasible solution. So we only move a fraction of the maximum possible distance. For example, with $r = 0.9$, we only go “90% of the way to the wall”. Thus our update is $\mathbf{1} + r(\beta d)$.

The transformed problem has now served its purpose of pushing the search direction away from the boundaries. We discard it and transform the solution \hat{x} back to the original LP to get the next iterate x^1 . As explained in Step 1, this is simply achieved through scaling by X^0 .