10 Karmarkars Algorithm

We want to solve the following linear program:

- $\min v = c^t x$ subject to $Ax = 0$ and $x \in \Delta$.
- Here $\Delta = \{x \in \mathbb{R}^n \mid e^t x = 1, x \geq 0\}$ with $e^t = (1, \dots, 1)$ denotes the standard simplex in $\mathbb{R}^n$.

Further assumptions:

1. A is an $m \times n$-matrix with rank m.
2. $Ae = 0$, i.e., the center of the simplex is feasible.
3. The optimum solution is 0.

10 Karmarkars Algorithm

Suppose you start with $\max\{c^t x \mid Ax = b; x \geq 0\}$.

- Multiply c by -1 and do a minimization. $\Rightarrow$ minimization problem
- We can check for feasibility by using the two phase algorithm. $\Rightarrow$ can assume that LP is feasible.
- Compute the dual; pack primal and dual into one LP and minimize the duality gap. $\Rightarrow$ optimum is 0
- Add a new variable pair x_ℓ, x'_ℓ (both restricted to be positive) and the constraint $\sum_i x_i = 1$. $\Rightarrow$ solution in simplex
- Add $-(\sum_i x_i) b_i = -b_i$ to every constraint. $\Rightarrow$ vector b is 0
- If A does not have full row rank we can delete constraints (or conclude that the LP is infeasible).
  $\Rightarrow$ A has full row rank

We still need to make e/n feasible.

10 Karmarkars Algorithm

The algorithm computes strictly feasible interior points $x^{(0)} = \frac{e}{n}, x^{(1)}, x^{(2)}, \dots$ with

$$c^t x^{(k)} \leq 2^{-\Theta(L)} c^t x^{(0)}$$

For $k = \Theta(L)$. A point x is strictly feasible if $x > 0$.

If my objective value is close enough to 0 (the optimum!!) I can "snap" to an optimum vertex.

10 Karmarkars Algorithm

Iteration:

1. Distort the problem by mapping the simplex onto itself so that the current point $\bar{x}$ moves to the center.
2. Project the optimization direction c onto the feasible region. Determine a distance to travel along this direction such that you do not leave the simplex (and you do not touch the border). $\hat{x}_{\text{new}}$ is the point you reached.
3. Do a backtransformation to transform $\hat{x}$ into your new point $\bar{x}_{\text{new}}$.

The Transformation

Let $\bar{Y} = \mathrm{diag}(\bar{x})$ the diagonal matrix with entries $\bar{x}$ on the diagonal.

Define
$$F_{\bar{x}} : x \mapsto \frac{\bar{Y}^{-1}x}{e^t\bar{Y}^{-1}x} \;.$$

The inverse function is
$$F_{\bar{x}}^{-1} : \hat{x} \mapsto \frac{\bar{Y}\hat{x}}{e^t\bar{Y}\hat{x}} \;.$$

Note that $\bar{x} > 0$ in every coordinate. Therefore the above is well defined.

Properties

$F_{\bar{x}}^{-1}$ really is the inverse of $F_{\bar{x}}$:

$$F_{\bar{x}}(F_{\bar{x}}^{-1}(\hat{x})) = \frac{\bar{Y}^{-1}\frac{\bar{Y}\hat{x}}{e^t\bar{Y}\hat{x}}}{e^t\bar{Y}^{-1}\frac{\bar{Y}\hat{x}}{e^t\bar{Y}\hat{x}}} = \frac{\hat{x}}{e^t\hat{x}} = \hat{x}$$

because $\hat{x} \in \Delta$.

Note that in particular every $\hat{x} \in \Delta$ has a preimage (Urbild) under $F_{\bar{x}}$.

Properties

$\bar{x}$ is mapped to e/n

$$F_{\bar{x}}(\bar{x}) = \frac{\bar{Y}^{-1}\bar{x}}{e^t\bar{Y}^{-1}\bar{x}} = \frac{e}{e^te} = \frac{e}{n}$$

Properties

A unit vectors e_i is mapped to itself:

$$F_{\bar{x}}(e_i) = \frac{\bar{Y}^{-1}e_i}{e^t\bar{Y}^{-1}e_i} = \frac{(0,\dots,0,1/\bar{x}_i,0,\dots,0)^t}{e^t(0,\dots,0,1/\bar{x}_i,0,\dots,0)^t} = e_i$$

Properties

All nodes of the simplex are mapped to the simplex:

$$F_{\bar{x}}(x) = \frac{\bar{Y}^{-1}x}{e^t\bar{Y}^{-1}x} = \frac{\left(\frac{x_1}{\bar{x}_1},\dots,\frac{x_n}{\bar{x}_n}\right)^t}{e^t\left(\frac{x_1}{\bar{x}_1},\dots,\frac{x_n}{\bar{x}_n}\right)^t} = \frac{\left(\frac{x_1}{\bar{x}_1},\dots,\frac{x_n}{\bar{x}_n}\right)^t}{\sum_i \frac{x_i}{\bar{x}_i}} \in \Delta$$

The Transformation

Easy to check:

- $F_{\bar{x}}^{-1}$ really is the inverse of $F_{\bar{x}}$.
- $\bar{x}$ is mapped to e/n.
- A unit vectors e_i is mapped to itself.
- All nodes of the simplex are mapped to the simplex.

10 Karmarkars Algorithm

We have the problem

$$\begin{aligned}
&\min\{c^t x \mid Ax = 0;\, x \in \Delta\}\\
&\quad= \min\{c^t F_{\bar{x}}^{-1}(\hat{x}) \mid AF_{\bar{x}}^{-1}(\hat{x}) = 0;\, F_{\bar{x}}^{-1}(\hat{x}) \in \Delta\}\\
&\quad= \min\{c^t F_{\bar{x}}^{-1}(\hat{x}) \mid AF_{\bar{x}}^{-1}(\hat{x}) = 0;\, \hat{x} \in \Delta\}\\
&\quad= \min\left\{\frac{c^t\bar{Y}\hat{x}}{e^t\bar{Y}\hat{x}} \mid \frac{A\bar{Y}\hat{x}}{e^t\bar{Y}\hat{x}} = 0;\, \hat{x} \in \Delta\right\}
\end{aligned}$$

Since the optimum solution is 0 this problem is the same as

$$\min\{\hat{c}^t\hat{x} \mid \hat{A}\hat{x} = 0, \hat{x} \in \Delta\}$$

with $\hat{c} = \bar{Y}^t c = \bar{Y}c$ and $\hat{A} = A\bar{Y}$.

Note that $e^t\bar{Y}x > 0$ for $x \in \Delta$.

We still need to make e/n feasible.

- We know that our LP is feasible. Let $\bar{x}$ be a feasible point.
- Apply $F_{\bar{x}}$, and solve

$$\min\{\hat{c}^t x \mid \hat{A}x = 0;\, x \in \Delta\}$$

- The feasible point is moved to the center.

10 Karmarkars Algorithm

When computing $\hat{x}_{\text{new}}$ we do not want to leave the simplex or touch its boundary (why?).

For this we compute the radius of a ball that completely lies in the simplex.

$$B\left(\frac{e}{n},\rho\right) = \left\{x \in \mathbb{R}^n \mid \left\|x - \frac{e}{n}\right\| \leq \rho\right\} .$$

We are looking for the largest radius r such that

$$B\left(\frac{e}{n},r\right) \cap \left\{x \mid e^t x = 1\right\} \subseteq \Delta.$$

10 Karmarkars Algorithm

This holds for $r = \|\frac{e}{n} - (e - e_1)\frac{1}{n-1}\|$. ($r$ is the distance between the center e/n and the center of the $(n-1)$-dimensional simplex obtained by intersecting a side ($x_i = 0$) of the unit cube with Δ.)

This gives $r = \frac{1}{\sqrt{n(n-1)}}$.

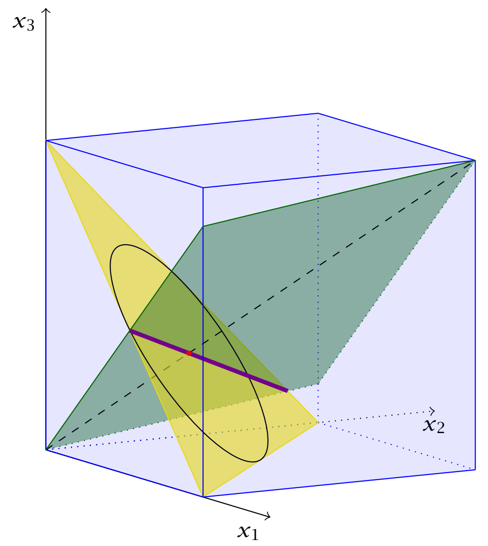
Now we consider the problem

$$\min\{\hat{c}^t x \mid \hat{A}x = 0, x \in B(e/n, r) \cap \Delta\}$$

This problem is easy to solve!!!

$$r^2 = (n-1)\cdot\left(\frac{1}{n} - \frac{1}{n-1}\right)^2 + \frac{1}{n^2} = \frac{1}{n^2(n-1)} + \frac{1}{n^2} = \frac{1}{n(n-1)}$$

The Simplex



10 Karmarkars Algorithm

Ideally we would like to go in direction of $-\hat{c}$ (starting from the center of the simplex).

However, doing this may violate constraints $\hat{A}\hat{x} = 0$ or the constraint $\hat{x} \in \Delta$.

Therefore we first project $\hat{c}$ on the nullspace of

$$B = \begin{pmatrix} \hat{A} \\ e^t \end{pmatrix}$$

We use

$$P = I - B^t(BB^t)^{-1}B$$

Then

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

is the required projection.

10 Karmarkars Algorithm

We get the new point

$$\hat{x}(\rho) = \frac{e}{n} - \rho\frac{\hat{d}}{\|\hat{d}\|}$$

for $\rho < r$.

Choose $\rho = \alpha r$ with $\alpha = 1/4$.

Iteration of Karmarkars Algorithm

- Current solution $\bar{x}$. $\bar{Y} := \mathrm{diag}(\bar{x}_1, \dots, \bar{x}_n)$.
- Transform problem via $F_{\bar{x}}(x) = \frac{\bar{Y}^{-1}x}{e^t\bar{Y}^{-1}x}$.
  Let $\hat{c} = \bar{Y}c$, and $\hat{A} = A\bar{Y}$.
- Compute
  $$\hat{d} = (I - B^t(BB^t)^{-1}B)\hat{c} \ ,$$
  where $B = \begin{pmatrix}\hat{A}\\ e^t\end{pmatrix}$.
- Set
  $$\hat{x}_{\text{new}} = \frac{e}{n} - \rho\frac{\hat{d}}{\|\hat{d}\|} \ ,$$
  with $\rho = \alpha r$ with $\alpha = 1/4$ and $r = 1/\sqrt{n(n-1)}$.
- Compute $\bar{x}_{\text{new}} = F_{\bar{x}}^{-1}(\hat{x}_{\text{new}})$.

Lemma 2
The new point $\hat{x}_{\text{new}}$ in the transformed space is the point that minimizes the cost $\hat{c}^t\hat{x}$ among all feasible points in $B(\frac{e}{n}, \rho)$.

Proof: Let $\hat{z}$ be another feasible point in $B(\frac{e}{n}, \rho)$.

As $\hat{A}\hat{z} = 0$, $\hat{A}\hat{x}_{\text{new}} = 0$, $e^t\hat{z} = 1$, $e^t\hat{x}_{\text{new}} = 1$ we have

$$B(\hat{x}_{\text{new}} - \hat{z}) = 0 \ .$$

Further,

$$\begin{aligned}(\hat{c} - \hat{d})^t &= (\hat{c} - P\hat{c})^t \\ &= (B^t(BB^t)^{-1}B\hat{c})^t \\ &= \hat{c}^tB^t(BB^t)^{-1}B\end{aligned}$$

Hence, we get

$$(\hat{c} - \hat{d})^t(\hat{x}_{\text{new}} - \hat{z}) = 0 \text{ or } \hat{c}^t(\hat{x}_{\text{new}} - \hat{z}) = \hat{d}^t(\hat{x}_{\text{new}} - \hat{z})$$

which means that the cost-difference between $\hat{x}_{\text{new}}$ and $\hat{z}$ is the same measured w.r.t. the cost-vector $\hat{c}$ or the projected cost-vector $\hat{d}$.

But

$$\frac{\hat{d}^t}{\|\hat{d}\|}(\hat{x}_{\text{new}} - \hat{z}) = \frac{\hat{d}^t}{\|\hat{d}\|}\left(\frac{e}{n} - \rho\frac{\hat{d}}{\|\hat{d}\|} - \hat{z}\right) = \frac{\hat{d}^t}{\|\hat{d}\|}\left(\frac{e}{n} - \hat{z}\right) - \rho < 0$$

as $\frac{e}{n} - \hat{z}$ is a vector of length at most ρ.

This gives $\hat{d}(\hat{x}_{\text{new}} - \hat{z}) \le 0$ and therefore $\hat{c}\hat{x}_{\text{new}} \le \hat{c}\hat{z}$.

In order to measure the progress of the algorithm we introduce a potential function f:

$$f(x) = \sum_j \ln(\frac{c^t x}{x_j}) = n\ln(c^t x) - \sum_j \ln(x_j) \ .$$

- The function f is invariant to scaling (i.e., $f(kx) = f(x)$).
- The potential function essentially measures cost (note the term $n\ln(c^t x)$) but it penalizes us for choosing x_j values very small (by the term $-\sum_j \ln(x_j)$; note that $-\ln(x_j)$ is always positive).

For a point $\hat{z}$ in the transformed space we use the potential function

$$\begin{aligned}\hat{f}(\hat{z}) &:= f(F_{\bar{x}}^{-1}(\hat{z})) = f(\frac{\bar{Y}\hat{z}}{e^t\bar{Y}\hat{z}}) = f(\bar{Y}\hat{z})\\ &= \sum_j \ln(\frac{c^t\bar{Y}\hat{z}}{\bar{x}_j\hat{z}_j}) = \sum_j \ln(\frac{\hat{c}^t\hat{z}}{\hat{z}_j}) - \sum_j \ln\bar{x}_j\end{aligned}$$

Observation:
This means the potential of a point in the transformed space is simply the potential of its pre-image under F.

Note that if we are interested in potential-change we can ignore the additive term above. Then f and $\hat{f}$ have the same form; only c is replaced by $\hat{c}$.

The basic idea is to show that one iteration of Karmarkar results in a constant decrease of $\hat{f}$. This means

$$\hat{f}(\hat{x}_{\text{new}}) \le \hat{f}(\frac{e}{n}) - \delta \ ,$$

where δ is a constant.

This gives

$$f(\bar{x}_{\text{new}}) \le f(\bar{x}) - \delta \ .$$

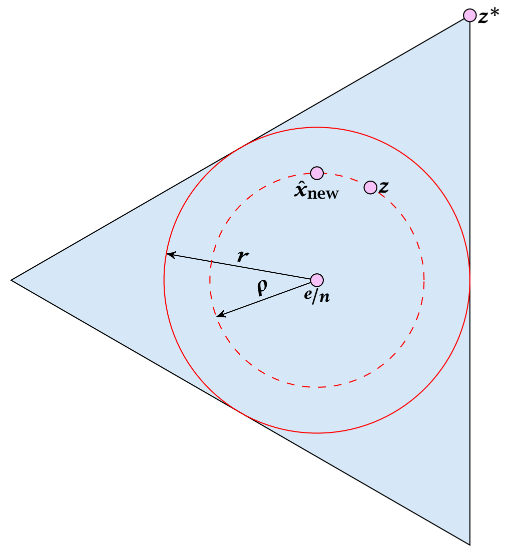
Lemma 3

There is a feasible point z (i.e., $\hat{A}z = 0$) in $B(\frac{e}{n}, \rho) \cap \Delta$ that has

$$\hat{f}(z) \le \hat{f}(\frac{e}{n}) - \delta$$

with $\delta = \ln(1 + \alpha)$.

Note that this shows the existence of a good point within the ball. In general it will be difficult to find this point.



Let z^* be the feasible point in the transformed space where $\hat{c}^t x$ is minimized. (Note that in contrast $\hat{x}_{\text{new}}$ is the point in the intersection of the feasible region and $B(\frac{e}{n}, \rho)$ that minimizes this function; in general $z^* \neq \hat{x}_{\text{new}}$)

z^* must lie at the boundary of the simplex. This means $z^* \notin B(\frac{e}{n}, \rho)$.

The point z we want to use lies farthest in the direction from $\frac{e}{n}$ to z^*, namely

$$z = (1 - \lambda)\frac{e}{n} + \lambda z^*$$

for some positive $\lambda < 1$.

Hence,

$$\hat{c}^t z = (1 - \lambda)\hat{c}^t \frac{e}{n} + \lambda \hat{c}^t z^*$$

The optimum cost (at z^*) is zero.

Therefore,

$$\frac{\hat{c}^t \frac{e}{n}}{\hat{c}^t z} = \frac{1}{1 - \lambda}$$

The improvement in the potential function is

$$
\begin{aligned}
\hat{f}(\tfrac{e}{n}) - \hat{f}(z) &= \sum_j \ln(\frac{\hat{c}^t \frac{e}{n}}{\frac{1}{n}}) - \sum_j \ln(\frac{\hat{c}^t z}{z_j}) \\
&= \sum_j \ln(\frac{\hat{c}^t \frac{e}{n}}{\hat{c}^t z} \cdot \frac{z_j}{\frac{1}{n}}) \\
&= \sum_j \ln(\frac{n}{1-\lambda} z_j) \\
&= \sum_j \ln(\frac{n}{1-\lambda}((1-\lambda)\frac{1}{n} + \lambda z_j^*)) \\
&= \sum_j \ln(1 + \frac{n\lambda}{1-\lambda} z_j^*)
\end{aligned}
$$

We can use the fact that for non-negative s_i

$$\sum_i \ln(1+s_i) \geq \ln(1+\sum_i s_i)$$

This gives

$$
\begin{aligned}
\hat{f}(\tfrac{e}{n}) - \hat{f}(z) &= \sum_j \ln(1 + \frac{n\lambda}{1-\lambda} z_j^*) \\
&\geq \ln(1 + \frac{n\lambda}{1-\lambda})
\end{aligned}
$$

Suppose true for $s_1, \dots, s_{k-1}$. Then

$$
\begin{aligned}
\sum_{i=1}^{k} \ln(1+s_i) &\geq \ln(1+\textstyle\sum_{i=1}^{k-1} s_i) + \ln(1+s_k) = \ln\Big((1+\textstyle\sum_{i=1}^{k-1} s_i)(1+s_k)\Big) \\
&= \ln\Big(1 + \sum_i s_i + s_k \textstyle\sum_{i=1}^{k-1} s_i\Big) \geq \ln(1+\textstyle\sum_i s_i)
\end{aligned}
$$

In order to get further we need a bound on λ:

$$\alpha r = \rho = \|z - e/n\| = \|\lambda(z^* - e/n)\| \leq \lambda R$$

Here R is the radius of the ball around $\frac{e}{n}$ that contains the whole simplex.

$R = \sqrt{(n-1)/n}$. Since $r = 1/\sqrt{(n-1)n}$ we have $R/r = n-1$ and

$$\lambda \geq \alpha \frac{r}{R} \geq \alpha/(n-1)$$

Then

$$1 + n\frac{\lambda}{1-\lambda} \geq 1 + \frac{n\alpha}{n-\alpha-1} \geq 1+\alpha$$

This gives the lemma.

Lemma 4
If we choose $\alpha = 1/4$ and $n \geq 4$ in Karmarkars algorithm the point $\hat{x}_{\text{new}}$ satisfies

$$\hat{f}(\hat{x}_{\text{new}}) \leq \hat{f}(\tfrac{e}{n}) - \delta$$

with $\delta = 1/10$.

Proof:

Define

$$g(\hat{x}) = n \ln \frac{\hat{c}^t \hat{x}}{\hat{c}^t \frac{e}{n}} = n(\ln \hat{c}^t \hat{x} - \ln \hat{c}^t \frac{e}{n}) \ .$$

This is the change in the cost part of the potential function when going from the center $\frac{e}{n}$ to the point $\hat{x}$ in the transformed space.

Similar, the penalty when going from $\frac{e}{n}$ to w increases by

$$h(\hat{x}) = \mathrm{pen}(\hat{x}) - \mathrm{pen}(\frac{e}{n}) = -\sum_j \ln \frac{\hat{x}_j}{\frac{1}{n}}$$

where $\mathrm{pen}(v) = -\sum_j \ln(v_j)$.

We want to derive a lower bound on

$$\begin{aligned} \hat{f}(\frac{e}{n}) - \hat{f}(\hat{x}_{\text{new}}) = & [\hat{f}(\frac{e}{n}) - \hat{f}(z)] \\ & + h(z) \\ & - h(\hat{x}_{\text{new}}) \\ & + [g(z) - g(\hat{x}_{\text{new}})] \end{aligned}$$

where z is the point in the ball where $\hat{f}$ achieves its minimum.

We have

$$[\hat{f}(\frac{e}{n}) - \hat{f}(z)] \geq \ln(1+\alpha)$$

by the previous lemma.

We have

$$[g(z) - g(\hat{x}_{\text{new}})] \geq 0$$

since $\hat{x}_{\text{new}}$ is the point with minimum cost in the ball, and g is monotonically increasing with cost.

We show that the change $h(w)$ in penalty when going from e/n to w fulfills

$$|h(w)| \le \frac{\beta^2}{2(1-\beta)}$$

where $\beta = n\alpha r$ and w is some point in the ball $B(\frac{e}{n}, \alpha r)$.

Hence,

$$\hat{f}(\frac{e}{n}) - \hat{f}(\hat{x}_{\text{new}}) \ge \ln(1+\alpha) - \frac{\beta^2}{(1-\beta)} \ .$$

Lemma 5

For $|x| \le \beta < 1$

$$|\ln(1+x) - x| \le \frac{x^2}{2(1-\beta)} \ .$$

For $|x| < 1$

$$\ln(1+x) = \sum_{i\ge1}(-1)^{i+1}\frac{x^i}{i} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

This gives

$$|\ln(1+x) - x| \le \left|-\frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots\right| \le \left|\frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \dots\right|$$

$$\le \frac{x^2}{2}\left|x^0 + x^1 + x^2 + \dots\right| = \frac{x^2}{2(1-|x|)} \ .$$

This gives for $w \in B(\frac{e}{n}, \rho)$

$$\begin{aligned}
|h(w)| &= \left|\sum_j \ln\frac{w_j}{1/n}\right| \\
&= \left|\sum_j \ln\left(\frac{1/n + (w_j - 1/n)}{1/n}\right) - \underbrace{\sum_j n\left(w_j - \frac{1}{n}\right)}_{=0}\right| \\
&= \left|\sum_j \left[\ln\left(1 + \underbrace{n(w_j - 1/n)}_{\le n\boxed{x} < 1}\right) - \underbrace{n(w_j - 1/n)}_{\boxed{x}}\right]\right| \\
&\le \sum_j \frac{n^2(w_j - 1/n)^2}{2(1-\alpha n r)} \\
&\le \frac{(\alpha n r)^2}{2(1-\alpha n r)}
\end{aligned}$$

The decrease in potential is therefore at least

$$\ln(1+\alpha) - \frac{\beta^2}{1-\beta}$$

with $\beta = n\alpha r = \alpha\sqrt{\frac{n}{n-1}}$.

It can be shown that this is at least $\frac{1}{10}$ for $n \ge 4$ and $\alpha = 1/4$.

Let $\bar{x}^{(k)}$ be the current point after the k-th iteration, and let $\bar{x}^{(0)} = \frac{e}{n}$.

Then $f(\bar{x}^{(k)}) \leq f(e/n) - k/10$.
This gives

$$
\begin{aligned}
n \ln \frac{c^t \bar{x}^{(k)}}{c^t \frac{e}{n}} &\leq \sum_j \ln \bar{x}_j^{(k)} - \sum_j \ln \frac{1}{n} - k/10 \\
&\leq n \ln n - k/10
\end{aligned}
$$

Choosing $k = 10n(\ell + \ln n)$ with $\ell = \Theta(L)$ we get

$$
\frac{c^t \bar{x}^{(k)}}{c^t \frac{e}{n}} \leq e^{-\ell} \leq 2^{-\ell} \ .
$$

Hence, $\Theta(nL)$ iterations are sufficient. One iteration can be performed in time $\mathcal{O}(n^3)$.