Mathematics 2: Interior point method

David Ocepek

August 11, 2022

1 Implementation

Our implementation closely followed the implementations proposed in the two papers given two us for this assignment with one clear exception being the determination of upper and lower bounds. For brevity we will focus on this part.

The proposed bounds in the papers is extremely inefficient, since they focus on the feasible region for x and not the feasible region for the solution x^*. As such we implemented our own heuristic which we explain in further detail later on in this section. First we give two examples from our test cases: the diet problem and the problem in Eq 1 where the solution x^* has non-integral coefficients

$$\begin{aligned}
&\min c^T x \\
&\text{s. t. } Ax \le b \\
&A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, b = [0, 0, 0.1, 0.5]^T, c = [-1, -1]^T
\end{aligned} \tag{1}$$

Heuristic: (Upper bound) Take i-th row $a_i x \le b_i$ if $\forall j : a_{ij} > 0$ then $x_j \le b/a_j$. It is equal when all other coefficients of x equal zero.

If some $a_{ij} = 0$ the corresponding $x_j \le \infty$ in other words it is not bounded.

If a single $a_{ij} < 0$ all coefficients are unbounded. However we can use bound from other rows where all $a_{ij} \ge 0$. We distinguish two cases:

Case 1: all x_j coresponding to negative a_{ij} are bounded by other rows, then $M = max_i a_{ij} x_j$ and $b_i^{'} = b_i - M$. We can then use our heuristic for non-negative a_i to get the bounds.

Case 2: not all x_j coresponding to negative a_{ij} are bounded by other rows then problem unbounded.

The lower bound is derived in a similar fashion. One notable difference is that the lower bound must be divided by our upper bound W since we normalize b with it. Furthermore Case 2 does not result in unboundedness (in fact this is the only case were bounds from the paper outperformed our bounds).

The comparison of upper and lower bounds between the paper and our heuristic are shown in table 1.

problem	upper bnd.	lower bnd.
diet(proposed)	3e40	5e-48
diet(actual)	2400	0.001
our test (proposed)	43750	1e-9
our test (actual)	0.5	-

Table 1: Inefficiency of proposed bounds

From table 1 we see that our heuristic does not always provide us with lower bounds. We performed 7 tests and the upper bound was always several orders of magnitude smaller than the proposed one, while the lower higher when it was present. We note that for the bound check this problem becomes much more pronounced, making bound checking infeasible with the paper bounds due to limited floating point precision.

2 Results

In table 2 we show the diet proposed by our interior point method. To verify that our method works properly we compared our results with scipy-optimize interior point method. Both methods returned the same result (a diet costing 148.83 EUR) and while this is in no way an exhaustive empirical analysis of the working of our program it does indicate that our model most likely works properly. From table 2 we can see that the optimal diet consists of

food	quantity [g]
bread	623
veg. oil	59
other	0

Nutrient	value
CH	299.04
PR	68.53
FT	90.0
EN	2200.0

Table 2: Optimal diet

only bread and vegetable oil.

The diet while optimal is not a diet we or the reader would most likely enjoy, indicating that more restrictions would be needed to get a proper diet. Clearly the interior point method needs carefully crafted inputs to return anything usable.

Out of the eight foods we had available the interior point method choose only two, this is inline with the theory that states that the interior point method will converge in a vertex.

We tested our two methods on six additional toy problems to further validate that we implemented our method properly. For brevity we do not go into any more detail besides stating that both methods got the same results[1].

3 Analytic center

The analytic center for our LP-problem is defined as

$$\begin{aligned} &\max_x \prod_{i \in I} s(x)_i \\ &\text{s. t. } Ax \leq b \\ &\qquad s(x) = b - Ax \end{aligned} \tag{2}$$

where $Ax \leq b$ is a system of m linear inequalities with n variables, $s(x)$ is the slack vector and $I \subseteq \{1, ..., n\}$ is the set of coordinates/indices, for $\exists x : s(x)_i > 0$. Intuitively the analytic center is the vector that maximizes our slack variables.

Problem 3

Let $\Phi = \{x; Ax \leq b\}$ be the set of all feasible solutions to problem defined in Eq. 2, then we want to prove that there exist an $x \in \Phi$ such that $s(x)_i > 0$ for all $i \in I$.

We start by proving that Φ is a convex set, something which will be key for our actual proof

Proof (The intersection of a finite number of half-planes Φ in $\mathbb{R}^n$ is a convex set): Given any two points $x_1, x_2 \in \Phi$, then if Φ is a convex set the following must hold

$$A(\theta x_1 + (1-\theta)x_2) \in \Phi, \forall \theta \in [0, 1] \tag{3}$$

Rewriting 3 the definition of Φ we get

$$\begin{aligned} &A(\theta x_1 + (1-\theta)x_2) \leq b \\ &A(\theta x_1 + (1-\theta)x_2) = \theta A x_1 + (1-\theta)A x_2 \leq \theta b + (1-\theta) b = b \end{aligned} \tag{4}$$

[1] An interested reader can find our test examples in our jupyter notebook.

□

Thus we have proven that Φ is a convex set (geometrically represented as a polyhedron). Next we choose two vertices $A, B \in \Phi$. We define sets A_+ and A_0, which represent the indices where $s(A) > 0$ and $s(A) = 0$. Using the same conventions we define B_+ and B_0. Note $A_+ \cap B_+ \neq \emptyset$. Now let's take the line AB, by convexity every point on line AB is in Φ. Now we show that any point T on line AB, where $T \neq (A \text{ or } B)$ has a set of indices T_+ where $s(T) > 0$ equal to $A_+ \cup B_+$.

Proof (Properties of point on line between two vertices of a convex set): Let us assume the opposite that $\exists T \in AB$, where $s(T)_i = 0$ for some $i \in A_+ \cup B_+$. Because Φ is convex T cannot be intersected by the plane a_i at point T. This leaves only the possibility that AB lies on a plane given by the row vector a_i, however this would mean both A and B lie on plane a_i and therefore $s(A)_i = 0$, $s(B)_i = 0$ and $i \notin A_+ \cup B_+$. Therefore we have a contradiction. □

Now lets define set X such that $X = \{x \in \Phi | x_+ \neq \emptyset\}$. By definition $I = \bigcup_{x \in X} x_+$. We can now use the following iterative method to construct a $T_{|X|}$ which will have $T_{|X|+} = I$. The method is as follows

Algorithm 1 Interior point construction

```
T is point between X_1 and X_2
for X_i ∈ X − {X_1, X_2} do
    T is point between T and X_i
end for
return T
```

Thus we have constructed a point T, with $T_+ = I$.

Problem 4

The analytic center in Eq. 2 can be rewritten into a strict convex optimization problem using the logarithm transform

$$
\begin{aligned}
\min_x \; & -\sum_{i\in I} \log s(x)_i \\
\text{s. t. } & Ax \leq b \\
& s(x) = b - Ax
\end{aligned} \tag{5}
$$

Proof (Strict convexity): Rewriting the optimization function from Eq. 5 we get

$$f(x) = -\sum_{i\in I} \log\left(b_i - \sum_{u=1}^{n} a_i^{(u)} x_u\right)$$

the Hessian of which is

$$\frac{\partial}{\partial x_j \partial x_k} = \sum_{i\in I} \frac{a_i^{(j)} a_i^{(k)}}{\left(b_i - \sum_u a_i^{(u)} x_u\right)^2} \tag{6}$$

To show that $f(x)$ is strictly convex it suffices to show that the Hessian is *positive definite (PD)*. Since the sum of *PD* matrices is PD it suffices to show that $[A_i]_{jk} = a_i^{(j)} a_i^{(k)}$ is *PD*, where $A^2 = \sum_i A_i$ is our original hessian. We observe that $[A_i] = {a_i}^T a_i$. By definition a matrix M is *PD* if $x^T M x > 0, \forall x \neq 0$, applying this to A_i we get

$$x^T a_i^T a_i x = (a_i x)^T a_i x = \left(\sum_u a_i^{(u)} x_u\right)^2 \geq 0 \tag{7}$$

Clearly A is *PD* if $\exists (i \in I) : (a_i x) \neq 0$. Therefore $f(x)$ is strictly convex if the columns of A are independent. □

Problem 5

Since $f(x)$ is convex, we only need to prove that convex functions have only a unique minimum

$f(x)$ convex $\implies$ $f(\cdot)$ has one unique minimum.

Proof (Uniqueness of minimum for convex function): Suppose f has two local minima at x_1 and x_2, then by convexity

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2), \quad \forall t \in [0,1] \tag{8}$$

since $f(x_1) = f(x_2)$ we get

$$\begin{aligned}
f(tx_1 + (1-t)x_2) & \\
&\leq t f(x_1) + (1-t) f(x_2) \\
&= t f(x_1) + (1-t) f(x_1) \\
&= f(x_1) \qquad (9)
\end{aligned}$$

which is a contradiction, thus we have proven that f can only have one minimum. $\square$

Problem 6

Given the problem

$$Ax \leq b, \text{ where } \quad A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ a & 1 \end{bmatrix} \text{ and } b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad (10)$$

, then its analytic center is

$$\max_{x_1, x_2} x_1 x_2 (1 - a x_1 - x_2) \qquad (11)$$

Applying the logarithm transform to Eq. 11 we get

$$\min_{x_1, x_2} -[\log x_2 + \log x_1 + log(1 - a x_1 - x_2)] \qquad (12)$$

Taking the derivative of our problem we get the system of equation

$$\begin{aligned}
\frac{1}{x_1} - \frac{a}{1 - a x_1 - x_2} &= 0 \\
\frac{1}{x_2} - \frac{1}{1 - a x_1 - x_2} &= 0 \qquad (13)
\end{aligned}$$

Solving said system we get

$$\begin{aligned}
&1 - a x_1 - x_2 - a x_1 = 0 \\
&1 - a x_1 - 2 x_2 = 0 \\
&a x_1 - x_2 = 0 \\
&x_2 = a x_1 \\
&1 - 3 a x_1 = 0 \\
&x_1 = \frac{1}{3a}, x_2 = \frac{1}{3} \qquad (14) \\
& \qquad (15)
\end{aligned}$$

Problem 7
Given the problem

$$Ax \leq b, \text{ where } \quad A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \text{ and } b = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \tag{16}$$

then its analytic center is

$$\max_{x_1,x_2} x_1 x_2 (1 - x_1 - x_2)^2 \tag{17}$$

Applying the logarithm transform to Eq. 17 we get

$$\min_{x_1,x_2} -[\log x_2 + \log x_1 + 2log(1 - x_1 - x_2)] \tag{18}$$

Taking the derivative of our problem we get the system of equation

$$\begin{aligned} \frac{1}{x_1} - \frac{2}{1 - x_1 - x_2} = 0 \\ \frac{1}{x_2} - \frac{2}{1 - x_1 - x_2} = 0 \end{aligned} \tag{19}$$

Solving said system we get

$$\begin{aligned} &1 - 3x_1 - x_2 = 0 \\ &1 - x_1 - 3x_2 = 0 \\ &x_1 = 1 - 3x_2 \\ &-2 + 8x_2 = \\ &x_1 = x_2 = \frac{1}{4} \end{aligned}$$