1 Duals and duals and duals

We are given the following linear program:

$$\begin{aligned} \max \mathbf{c}^\mathsf{T}\mathbf{x}&, \\ \mathbf{A}\mathbf{x} \leq \mathbf{b}&, \\ \mathbf{x} \geq \mathbf{0}&. \end{aligned} \tag{P}$$

1. To find the dual problem, we look for an upper bound on cost. We multiply the constraints with vector $\mathbf{y}$ from the left. We get

   $$\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{x} \leq \mathbf{y}^\mathsf{T}\mathbf{b}.$$

   To keep the inequality direction, we need $\mathbf{y} \geq 0$, which are the first constraints of the dual. We want the obtained expression to be an upper bound on the cost:

   $$\mathbf{c}^\mathsf{T}\mathbf{x} \leq \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{x} \leq \mathbf{y}^\mathsf{T}\mathbf{b},$$

   meaning we need to solve

   $$\min \mathbf{b}^\mathsf{T}\mathbf{y}.$$

   This is the dual cost function. Because all elements of $\mathbf{x}$ are positive, we can also use $\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{x} = (\mathbf{A}^\mathsf{T}\mathbf{y})^\mathsf{T}\mathbf{x}$ to obtain constraints

   $$\mathbf{A}^\mathsf{T}\mathbf{y} \geq \mathbf{c}.$$

   All together, the dual problem is

   $$\begin{aligned} \min \mathbf{b}^\mathsf{T}\mathbf{y}&, \\ \mathbf{A}^\mathsf{T}\mathbf{y} \geq \mathbf{c}&, \\ \mathbf{y} \geq \mathbf{0}&. \end{aligned} \tag{D}$$

2. Now, we show that the linear program

   $$\begin{aligned} \max \mathbf{b}^\mathsf{T}\mathbf{y}&, \\ \mathbf{A}^\mathsf{T}\mathbf{y} + \mathbf{s} = \mathbf{c}&, \\ \mathbf{s} \geq \mathbf{0}&, \\ \mathbf{y} \in \mathbb{R}&, \end{aligned} \tag{D}$$

   from [1] is the dual of the primal problem (P):

   $$\begin{aligned} \min \mathbf{c}^\mathsf{T}\mathbf{x}&, \\ \mathbf{A}\mathbf{x} \geq \mathbf{b}&, \\ \mathbf{x} \geq \mathbf{0}&. \end{aligned} \tag{P}$$

   First, we factor out $\mathbf{s}$ and multiply it with $\mathbf{x}$ from the left:

   $$\mathbf{x}^\mathsf{T}\mathbf{s} = \mathbf{x}^\mathsf{T}(\mathbf{c} - \mathbf{A}^\mathsf{T}\mathbf{y}) = \mathbf{x}^\mathsf{T}\mathbf{c} - \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{y} = \mathbf{x}^\mathsf{T}\mathbf{c} - (\mathbf{A}\mathbf{x})^\mathsf{T}\mathbf{y}.$$

   We define the primal constraint $\mathbf{x} \geq \mathbf{0}$ which makes the upper dot product non-negative. We also define the constraints $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ and replace $\mathbf{A}\mathbf{x}$ in the previous line with something smaller, i.e., $\mathbf{b}$, which makes the value bigger. We obtain the inequalities

$$0 \leq \mathbf{x}^\mathsf{T}\mathbf{s} \leq \mathbf{c}^\mathsf{T}\mathbf{x} - \mathbf{b}^\mathsf{T}\mathbf{y}.$$

The only thing that remains is to show that maximising $\mathbf{b}^\mathsf{T}\mathbf{y}$ is the same as minimizing $\mathbf{c}^\mathsf{T}\mathbf{x}$. The obtained inequality tells us that the difference between these two dot products is non-negative. Thus,

$$\mathbf{b}^\mathsf{T}\mathbf{y} \leq \mathbf{c}^\mathsf{T}\mathbf{x}.$$

This shows that the minimum of the right-hand side is the maximum of the left side.

2 Bounding the unbounded

3. To prove Proposition 3, we first approximate the value of a determinant. Its definition is

$$\det \mathbf{A} = \sum_{\sigma} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)},$$

where σ are all permutations of length n. We can omit the sign to bound the absolute value. We know that there are $n!$ permutations, so we can compute an upper bound without the sum. Lastly, we use $|a_{ij}| \leq U$ to obtain

$$|\det \mathbf{A}| \leq n! U \cdots U = n! U^n \leq n^n U^n = (nU)^n.$$

Cramer's rule can be used to compute the solutions of a linear system. For its use, the matrix of the system has to be invertible, meaning that a unique solution exists. But, a system can have a solution even if the matrix is rectangular. If it has more rows than columns, the additional rows just need to agree with the solution obtained by a non-singular square submatrix, i.e., the additional hyperplanes need to intersect in the obtained solution. Similarly, if the system has more columns, we can take the solution of the non-singular subsystem and set the remaining variables to 0.

Because the proposition assumes the linear program is feasible, we can use the described idea. Using Cramer's rule, we can obtain the inequality:

$$|x_i| = \frac{|\det \mathbf{A}_i|}{|\det \mathbf{A}|} \leq |\det \mathbf{A}_i| \leq (nU)^n = M.$$

We can omit the denominator because the matrix elements are integers and the chosen submatrix is invertible, thus $1 \leq |\det \mathbf{A}|$. We define $(nU)^n = M$ and because $\mathbf{x} \geq \mathbf{0}$, we get

$$\mathbf{x} \leq M.$$

This shows that a feasible solution has to be bounded by M and proves the first part of the proposition.

The second part talks about the optimal solution. If we assume that the linear program has an optimal solution, it has to be feasible. Because all feasible solutions are bounded, we can also conclude that an optimal solution will be bounded.

■

3 Interior-point algorithm

We are given the following linear program:

$$\begin{aligned} \min -3x_1 - 4x_2&, \\ 3x_1 + 3x_2 + 3x_3 &= 4, \\ 3x_1 + x_2 + x_4 &= 3, \\ x_1 + 4x_2 + x_5 &= 4, \\ x_1, x_2, x_3, x_4, x_5 &\geq 0. \end{aligned} \tag{problem X}$$

5. From the definition of the linear program, we can obtain:

$$\mathbf{A} = \begin{bmatrix} 3 & 3 & 3 & 0 & 0 \\ 3 & 1 & 0 & 1 & 0 \\ 1 & 4 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 3 \\ 4 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} -3 \\ -4 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

   Using these, we can write the dual problem presented in the lectures and the paper:

$$\begin{aligned} \max 4y_1 + 3y_2 + 4y_3&, \\ 3y_1 + 3y_2 + y_3 + s_1 &= -3, \\ 3y_1 + y_2 + 4y_3 + s_2 &= -4, \\ 3y_1 + s_3 &= 0, \\ y_2 + s_4 &= 0, \\ y_3 + s_5 &= 0, \\ y_1, y_2, y_3 &\in \mathbb{R}, \\ s_1, s_2, s_3, s_4, s_5 &\geq 0. \end{aligned}$$

6. To test whether the vectors

$$\mathbf{x} \left[\frac{2}{5}, \frac{8}{15}, \frac{2}{5}, \frac{19}{15}, \frac{22}{15}\right]^\mathsf{T}, \qquad \mathbf{y} = \left[-\frac{4}{5}, -\frac{4}{5}, -\frac{2}{3}\right]^\mathsf{T}, \qquad \mathbf{s} = \left[\frac{37}{15}, \frac{28}{15}, \frac{12}{5}, \frac{4}{5}, \frac{2}{3}\right]^\mathsf{T},$$

   are feasible solutions to both problems, we plug them into the definition and check whether the right- and left-hand sides are equal. Using Python, we confirm that they are indeed feasible solutions. Since all elements of $\mathbf{x}$ and $\mathbf{s}$ are non-zero, the solutions are strictly feasible, i.e., they are strictly inside the feasibility region.

7. Because we have shown that the above vectors are strictly feasible solutions, invariants 1 and 2 from [1] hold. For the invariant 3, we need

$$\sigma^2 = \sum_{i=1}^{m} \left(\frac{x_i s_i}{\mu} - 1\right)^2 \leq \frac{1}{4}.$$

   Using the average of $x_i s_i$ for μ, we get $\sigma^2 = 0.002$, so the invariant holds. If we compute

$$\mathbf{x} \odot \mathbf{s} = \left[\frac{222}{225}, \frac{224}{225}, \frac{216}{225}, \frac{228}{225}, \frac{220}{225}\right]^\mathsf{T},$$

   we can see that all elements are close to 1, so $\mu = 1$ would work as well.

9. We stop the interior-point method iterations if the new solution is less than $1\mathrm{e}^{-11}$ away from the current solution. We obtain the minimum cost function of -4.9 with $\mu = 4\mathrm{e}^{-11}$ and

$$
\mathbf{x} = [0.44, 0.89, 0, 0.78, 0]^\mathsf{T}, \quad \mathbf{y} = [-0.89, 0, -0.33]^\mathsf{T}, \quad \mathbf{s} = [0, 0, 2.67, 0, 0.33]^\mathsf{T}.
$$

The zero elements of $\mathbf{x}$ and $\mathbf{s}$ were obtained by rounding the solutions to 5 decimal points, even though some numbers were already 0 before rounding. We obtained 5 additional equations:

$$
\begin{aligned}
x_3 &= 0, \\
x_5 &= 0, \\
s_1 &= 0, \\
s_2 &= 0, \\
s_4 &= 0.
\end{aligned}
$$

10. Using these additional equations, we can compute the exact solution of

$$
\begin{aligned}
\mathbf{A}\mathbf{x} &= \mathbf{b}, \\
\mathbf{A}^\mathsf{T}\mathbf{y} + \mathbf{s} &= \mathbf{c}.
\end{aligned}
$$

We create a new linear system $\mathbf{B}\mathbf{z} = \mathbf{d}$ by adding the new equations. All variables are now combined into $\mathbf{z} = [\mathbf{x}, \mathbf{y}, \mathbf{s}]^\mathsf{T}$. The system is defined as:

$$
\mathbf{B} = \left[\begin{array}{c|c|c}
\mathbf{A} & & \\ \hline
 & \mathbf{A}^\mathsf{T} & \mathbf{I} \\ \hline
\begin{matrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ & & & & \\ & & & & \\ & & & & \end{matrix} & & \begin{matrix} & & & & \\ & & & & \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{matrix}
\end{array}\right], \quad
\mathbf{d} = \left[\begin{array}{c}
\mathbf{b} \\ \hline
\mathbf{c} \\ \hline
0 \\ 0 \\ 0 \\ 0 \\ 0
\end{array}\right],
$$

where all other elements of $\mathbf{B}$ are set to 0. This system can be solved using a linear system solver and the obtained solution is the same as the vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{s}$, obtained using interior-point method. Also, the objective gap $\mathbf{c}^\mathsf{T}\mathbf{x} - \mathbf{b}^\mathsf{T}\mathbf{y}$ is 0. This shows that the solutions are feasible and optimal.

4 Commercial solver

12. As a commercial solver, we use `scipy`'s implementation of the simplex and interior-point methods. In both cases, the obtained solutions equal the ones from our implementation. Consequently, the obtained minimum cost value is the same one as well. Both methods find the solution only in a few iterations, 3 for the simplex and 5 for the interior-point method. This is the largest difference with our implementation, which takes more than 200 iterations for a comparable accuracy.

References

[1] Kurt Mehlhorn and Sanjeev Saxena. "A Still Simpler Way of Introducing the Interior-Point Method for Linear Programming". In: *CoRR* abs/1510.03339 (2015). arXiv: `1510.03339`. URL: `http://arxiv.org/abs/1510.03339`.