Bioinformatics I

WS 15/16

Tutor: Alexander Seitz

& Benjamin Schroeder

Assignment 5

(Handed in 16. November 2015)

Theoretical Assignment - Comparison with at most l mismatches

General we are now only looking at the worst case scenario, which means uniform distributed mistmatches in an alignment. Also we only look at k >= 1, as k = 0 is in our application not a useful result. Assume two sequences of length t with l mismatches.

Than both sequences contain l tuples of length $\lfloor \frac{t}{l+1} \rfloor$ and one tuple which has a length of maximal $\lfloor \frac{t}{l+1} \rfloor$. Where $k = \lfloor \frac{t}{l+1} \rfloor$ is the maximal tuple length, possible in both sequences.

So both sequences share l+1 k-tuples and for each $k \leq \lfloor \frac{t}{l+1} \rfloor$ they share $(l+1) * \lfloor \frac{\lfloor \frac{t}{l+1} \rfloor}{k} \rfloor$ k-tuples.

Theoretical Assignment - Linear programming by hand

The feasible region of this linear program is shown in figure 1, where the red line is constraint 1, green line is constraint 2, constraint 3 is drawn as a blue line and the yellow line is constraint 4.

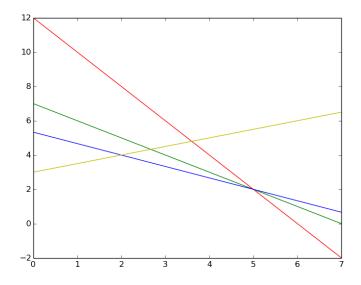


Figure 1: feasible region of linear program

Theoretical Assignment - Bonus: Carillo-Lipman bound