Column generation

Column generation or **delayed column generation** is an efficient algorithm for solving larger linear programs.

The overarching idea is that many linear programs are too large to consider all the variables explicitly. Since most of the variables will be non-basic and assume a value of zero in the optimal solution, only a subset of variables need to be considered in theory when solving the problem. Column generation leverages this idea to generate only the variables which have the potential to improve the objective function—that is, to find variables with negative reduced cost (assuming without loss of generality that the problem is a minimization problem).

The problem being solved is split into two problems: the master problem and the subproblem. The master problem is the original problem with only a subset of variables being considered. The subproblem is a new problem created to identify a new variable. The objective function of the subproblem is the reduced cost of the new variable with respect to the current dual variables, and the constraints require that the variable obey the naturally occurring constraints.

The process works as follows. The master problem is solved—from this solution, we are able to obtain dual prices for each of the constraints in the master problem. This information is then utilized in the objective function of the subproblem. The subproblem is solved. If the objective value of the subproblem is negative, a variable with negative reduced cost has been identified. This variable is then added to the master problem, and the master problem is re-solved. Re-solving the master problem will generate a new set of dual values, and the process is repeated until no negative reduced cost variables are identified. The subproblem returns a solution with non-negative reduced cost, we can conclude that the solution to the master problem is optimal.

In many cases, this allows large linear programs that had been previously considered intractable to be solved. The classical example of a problem where this is successfully used is the cutting stock problem. One particular technique in linear programming which uses this kind of approach is the Dantzig–Wolfe decomposition algorithm. Additionally, column generation has been applied to many problems such as crew scheduling, vehicle routing, and the capacitated p-median problem.