COLLAPSED DUAL SOLUTION TO LARGE MARKETING OPTIMIZATION PROBLEM

Bob Agnew, raagnew1@gmail.com, raagnew.com

We illustrate this method in the context of credit card campaign optimization, but it can easily be developed for any generalized assignment problem with linear objective function and linear constraints, including that a prospect can receive at most one of several offers. Our method relaxes the large binary assignment problem to a large primal linear program which has a large but simple dual linear program which in turn can be collapsed to a nonlinear form with few variables. This collapsed dual function is readily solved with good precision by a conventional solver. The dual solution is then employed to solve the large primal problem, again with good precision.

First a review of linear programming duality. We have a conventional setup with

Primal Linear Program	Dual Linear Program
Maximize: $c^T x$ Subject to: $Ax \le b$ $x \ge 0$	Minimize: $b^T y$ Subject to: $A^T y \ge c$ $y \ge 0$

where matrix A and column vectors b and c are specified while column vectors x and y represent solutions and T denotes transposition. For exact optimal solutions x^* and y^* we have $c^T x^* = b^T y^*$. More importantly for any feasible solutions x and y to the primal and dual programs respectively, we have $c^T x \le b^T y$. Hence, a feasible dual value always bounds any feasible primal value, including the exact optimal primal value, and the difference $b^T y - c^T x$ indicates algorithmic precision.

Now consider the following credit card campaign problem. p_{ij} is the estimated probability of a favorable response if prospect i gets offer j. Normally the number of prospects is huge and the number of alternative offers is small. We want to maximize the expected number of total responses, but there are constraints. First, offers have costs and there is an overall dollar budget upper bound. Second, prospects have (inverse) risk scores and there is an overall average risk score lower bound required for estimated responders, even after trimming extremely risky prospects. Finally, there is the constraint that a prospect can receive at most one offer. This might seem restrictive but it really isn't because a combo can be a separate offer and it has to have its own probability scores anyway. Mathematically we have

$$\text{Maximize: } \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij}\, x_{ij}$$

$$\text{Subject to: } \sum_{i=1}^{m} \sum_{j=1}^{n} c_j\, x_{ij} \;\leq\; b$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} \left(\bar{s} - s_i \right) x_{ij} \;\leq\; 0$$

$$\sum_{j=1}^{n} x_{ij} \;\leq\; 1 \text{ for all } i$$

$$x_{ij} \in \left\{ 0,1 \right\} \text{ for all } i \text{ and } j$$

Here c_j is the unit dollar cost of offer j, s_i is the risk score for prospect i (higher the better), b is the dollar budget upper bound, and $\bar{s}$ is the average responder risk score lower bound. x_{ij} is the binary assignment variable which equals 1 for an offer and 0 for no offer.

This generalized assignment problem can be relaxed to a continuous linear program by replacing $x_{ij} \in \left\{ 0,1 \right\}$ with $x_{ij} \geq 0$ for all i and j. We can then visualize this primal linear program in matrix form as

$$\text{Maximize: } \begin{pmatrix} p_{11} & \cdots & p_{1n} & \cdots & p_{m1} & \cdots & p_{mn} \end{pmatrix} \begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{mn} \end{pmatrix}$$

$$\text{Subject to: } \begin{pmatrix} c_1 & \cdots & c_n & \cdots & c_1 & \cdots & c_n \\ p_{11}(\bar{s}-s_1) & \cdots & p_{1n}(\bar{s}-s_1) & \cdots & p_{m1}(\bar{s}-s_m) & \cdots & p_{mn}(\bar{s}-s_m) \\ 1 & \cdots & 1 & 0 & \cdots & \cdots & 0 \\ & & & \ddots & & & \\ 0 & \cdots & \cdots & 0 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{mn} \end{pmatrix} \le \begin{pmatrix} b \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{mn} \end{pmatrix} \ge \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

The dual of this linear program can then be formulated as

$$\text{Minimize: } \begin{pmatrix} b & 0 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ z_1 \\ \vdots \\ z_m \end{pmatrix}$$

$$\text{Subject to: } \begin{pmatrix} c_1 & p_{11}(\bar{s}-s_1) & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ c_n & p_{1n}(\bar{s}-s_1) & 1 & \cdots & \vdots \\ \vdots & \vdots & 0 & \cdots & 0 \\ c_1 & p_{m1}(\bar{s}-s_m) & \vdots & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ c_n & p_{mn}(\bar{s}-s_m) & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ z_1 \\ \vdots \\ z_m \end{pmatrix} \geq \begin{pmatrix} p_{11} \\ \vdots \\ p_{1n} \\ \vdots \\ p_{m1} \\ \vdots \\ p_{mn} \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ z_1 \\ \vdots \\ z_m \end{pmatrix} \geq \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where we have labeled elements of the dual vector for convenience. This boils down to

$$\text{Minimize: } b\,y_1 + \sum_{i=1}^{m} z_i$$

$$\text{Subject to: } c_j\,y_1 + p_{ij}(\bar{s}-s_i)\,y_2 + z_i \geq p_{ij} \quad \text{for all } i \text{ and } j$$

$$y_1,\ y_2,\ z_1 \cdots z_m \geq 0$$

but clearly we must have $z_i = max\Big(0,\ max_j\big(\,p_{ij} - c_j\,y_1 - p_{ij}(\bar{s}-s_i)\,y_2\,\big)\Big)$ at a minimum and hence the dual problem can be collapsed to the two-variable problem

$$\text{Minimize: } b\, y_1 + \sum_{i=1}^{m} max\Big(0,\, max_j \big(p_{i\,j} - c_j\, y_1 - p_{i\,j} \left(\bar{s} - s_i \right) y_2 \big) \Big)$$

$$\text{Subject to: } y_1,\, y_2 \ \geq\ 0$$

This problem is not smooth, but it is continuous and convex. Moreover, it can be approximated nicely with a conventional solver. Given a quasi-optimal dual solution $y_1^*,\, y_2^*$, we define $d_{i\,j} = p_{i\,j} - c_j\, y_1^* - p_{i\,j} \left(\bar{s} - s_i \right) y_2^*$ and $w_i = max_j\, d_{i\,j}$. We then rank w_i descending and choose $x_{i\,j} = 1$ for the first j with $w_i = d_{i\,j}$ as long as $w_i \geq 0$ and cumulative cost is less than b . Thereafter $x_{i\,j} = 0$ for all j , i.e., prospect gets no offer.

The resulting solution is invariably sharp, but approximation can cause the average responder risk score to be slightly below the lower bound and the primal value may actually exceed the dual value by a tad. If the risk score bound absolutely has to be met, then the lower bound can be tweaked accordingly. Computation time to run various scenarios is not an issue. If some other expected value score $v_{i\,j}$ is available, it can be substituted for the probability score $p_{i\,j}$ in the primal objective function, but not in the constraints, and the math works through in the same way.