Dual Solution of a Large Marketing Optimization Problem

Consider the following optimization problem which is frequently encountered in direct marketing applications. We have a population of n prospects and m offers, where n is usually quite large and m is greater than one but relatively small. If prospect j receives offer i, expected profit is p_{ij}, where "profit" can be any criterion which is additive across prospects and offers. Our objective is to maximize total expected profit, subject to a quantity constraint for each offer and the requirement that a prospect can receive *at most* one offer.

Mathematically, we have:

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

$$\text{subject to} \quad \sum_{j=1}^{n} x_{ij} = a_i \; ; \; i = 1,\dots,m$$

$$\sum_{i=1}^{m} x_{ij} \le 1 \; ; \; j = 1,\dots,n$$

$$x_{ij} \in \{0,1\} \; ; \; i = 1,\dots,m \; , \; j = 1,\dots,n$$

where a_i is the required total integer quantity for offer i and binary decision variable x_{ij} equals one if prospect j receives offer i, zero otherwise.

Structurally, this problem is very simple. In fact, it can be viewed as a transportation problem or a maximum generalized assignment problem, so well-known efficient algorithms are available for a problem of limited size. What distinguishes direct marketing problems is their very large prospect dimension, typically in the millions. Our approach to approximating the optimal solution for such a large problem represents a straightforward application of duality which should be familiar to OR researchers, practitioners, and students per Sokol (2007). More general constraints could be added, but we keep the problem simple for illustration.

We note that at least three commercial programs are designed to approximate optimal solutions to large-scale direct marketing problems with a variety of constraints: Marketswitch Optimization; SAS Marketing Optimization; and Unica Affinium Campaign Optimize. These programs incorporate a variety of algorithmic approaches to approximate optimal solutions and they all work well. Our discussion is not intended to detract from commercially available software, only to share some insights about a dual-driven approach to large direct marketing optimization problems.

Following a familiar heuristic path, we relax the problem above by substituting $x_{ij} \geq 0$ for the binary conditions, thereby obtaining a linear program. Its dual LP has the form:

$$\text{minimize} \quad \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} v_j$$

$$\text{subject to} \quad u_i + v_j \geq p_{ij} \; ; \; i = 1,\ldots,m \; , \; j = 1,\ldots,n$$

$$v_j \geq 0 \; ; \; j = 1,\ldots,n$$

where the unrestricted u_i are associated with the m offer quantity constraints and the nonnegative v_j are associated with the n prospect bounds. Clearly, $v_j \geq p_{ij} - u_i$ for all i so that minimization implies $v_j = max(0, max_i(p_{ij} - u_i))$. It follows that the dual can be collapsed into the unconstrained problem:

$$\text{minimize} \quad \sum_{i=1}^{m} a_i u_i + \sum_{j=1}^{n} max(0, max_i(p_{ij} - u_i)) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_i u_i / n + max(0, max_i(p_{ij} - u_i)) \right) .$$

This problem involves an extensive summation across prospects, but it has a limited number of variables. Moreover, it is continuous and convex. It is not smooth, but it turns out that ordinary nonlinear solvers with estimated derivatives work quite nicely on this problem, i.e., it isn't necessary to employ a subgradient or other algorithm designed expressly for a nonsmooth objective function. For problems with limited prospect populations (i.e., up to 10,000), even the Excel standard solver will do. For larger problems, we have employed PROC NLP in SAS/OR with the Quasi-Newton option. This general nonlinear programming procedure has a nice internal data step which automatically sums the objective function at each iteration. For very large problems (i.e., tens of millions of prospects), a workable dual solution can be generated on a random sample of prospects, with the prospect summation across the sample in the collapsed dual objective function but keeping divisor n equal to total prospect population size.

Given an approximately optimal dual solution $u_1,...,u_m$, the rest is pretty simple. Letting $d_{ij} = p_{ij} - u_i$, we put $v_j = max_i\, d_{ij}$ and sort the entire prospect file descending by v_j . In accordance with dual complementarity, we then go down the file and assign each prospect j to an offer i with maximum d_{ij} across all offers not already fully allocated. This is exceedingly easy and efficient to do in SAS.

Offer	Quantity	Dual Variable	Expected Profit
1	300,000	44.03	21,767,810
2	200,000	31.96	9,270,203
3	100,000	26.49	3,352,259
None	400,000		
Total	1,000,000		34,390,272
Dual Objective Function			34,390,313

Table 1. Illustrative Optimization Results

Table 1 contains illustrative results for a problem with 1,000,000 prospects and three offers with required quantities 300,000, 200,000, and 100,000 respectively, which of course means that 400,000 prospects receive no offer. The profits for this problem are simulated, but the results are indicative of numerous real-world applications of this method in large-scale direct marketing campaigns at a leading financial services firm. Specifically, we put $p_{1j} = 100\, z_{1j} - 10$, $p_{2j} = .6\, p_{1j} + .4\left(60\, z_{2j} - 6\right)$, and $p_{3j} = .4\, p_{1j} + .6\left(40\, z_{3j} - 4\right)$, where the z_{ij} are pseudo-random numbers in $(0,1)$, thus allowing for some typical correlation of offer profits across prospects. Note that the dual objective function provides an upper bound on achievable total expected profit and that we are very close. Note also that the offer dual variables have their usual shadow price interpretation with respect to constraint relaxation, although we are simply using them for algorithmic purposes. We ran this problem on an IBM P595 AIX shared server. The entire run took about 5 minutes of CPU time, most of which involved dual optimization. The results speak for themselves. Our approach achieves excellent precision with efficient use of computer resources.

References

Marketswitch Optimization. http://www.marketswitch.com

SAS Marketing Optimization. http://www.sas.com/solutions/crm/mktopt

SAS/OR 9.2 User's Guide: Mathematical Programming. 2008.
http://support.sas.com/documentation/cdl/en/ormpug/59679/PDF/default/ormpug.pdf

Sokol, J. 2007. Duality is So Useful! *OR/MS Today* **34** (6) 10.
http://www.lionhrtpub.com/orms/orms-12-07/freducation.html

Unica Affinium Campaign Optimize.
http://www.unica.com/products/campaign_management.htm#optimize

Robert A. Agnew
raagnew@comcast.net
http://home.comcast.net/~raagnew