Continuous Equality Knapsack with Probit-Style Objectives*

Jamie Fravel[1], Robert Hildebrand[†1], and Laurel Travis[2]

[1]Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, Virginia, USA
[2]Business Information Technology, Virginia Tech, Blacksburg, Virginia, USA

Abstract

We study continuous, equality knapsack problems with uniform separable, non-convex objective functions that are continuous, strictly increasing, antisymmetric about a point, and have concave and convex regions. For example, this model captures a simple allocation problem with the goal of optimizing an expected value where the objective is a sum of cumulative distribution functions of identically distributed normal distributions (i.e., a sum of inverse probit functions). We prove structural results of this model under general assumptions and provide two algorithms for efficient optimization: (1) running in linear time and (2) running in a constant number of operations given preprocessing of the objective function.

1 Introduction

We study the following restricted version of a nonlinear continuous knapsack problem

Model 1 (x-space Non-linear Knapsack)**.**

$$\max \left\{ F(\mathbf{x}) = \sum_{i=1}^{n} f(x_i) : \sum_{i=1}^{n} x_i = M, \mathbf{x} \in [a,b]^n \right\}. \tag{1}$$

where $f\colon [a,b] \to \mathbb{R}$, and $a, b, M \in \mathbb{R}$ with $a \leq b \leq M$.

We make use of the following assumptions on f for our main results: (A1) f is continuous and strictly increasing, (A2) f antisymmetric about a point $c \in (a,b)$ (i.e. $f(x) - f(c) = f(c) - f(c - (x - c))$ for all $x \in [a,b]$), and (A3) f is twice differentiable such that $f''(x) < 0$ for all $x \in (c,b)$. By (A2) and (A3) we also have $f''(x) > 0$ for all $x \in (a,c)$. In particular, f is strictly convex on $[a,c]$ and strictly concave on $[c,b]$. See Figure 1 for an example.

Several classes of functions satisfy these assumptions including logistic models, probit models, arc-tangent functions, and many cumulative distribution functions. For example, these results connect to analysis of racial representation in legislative redistricting in the United States by using

$$f(r) := \Phi\left(\beta \cdot r - \beta_0\right), \tag{2}$$

*J. Fravel and R. Hildebrand were partially funded by ONR Grant N00014-20-1-2156. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research. R. Hildebrand was also partially funded by the 2020-2021 Policy+ grant from Virginia Tech.

[†]Corresponding author: rhil@vt.edu

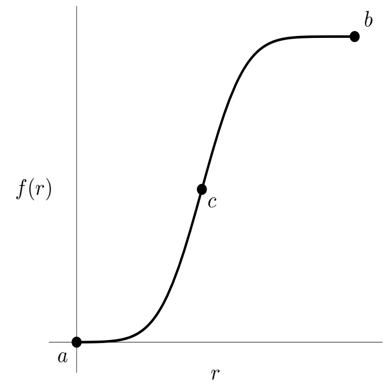


Figure 1: A function which satisfies our assumptions (A1), (A2) and (A3).

where Φ is the cumulative distribution function of the standard normal distribution and $r = \frac{\text{BVAP}}{\text{VAP}}$, that is, the ratio of black voters to the total population of voters in a given district. After calibrating β and β_0 from historical data, this function predicts the probability that a district will elect a black congressional representative. More details on this work can be seen in [5, 6, 7][1].

Prior work Knapsack and resource allocation problems are well studied in the literature, particularly for linear objectives. We mention just a few references that are most related to our work. Perhaps the earliest reference with nonlinear objectives is Derman[4] in 1959. Derman studies a similar variant to our allocation problem and then handles allocation under random demand. Derman does not require separability of the objective function. A similar setup was used under an environmental application in [8] to identify a starting point for a simulation optimization heuristic with the objective of maximizing the probability that an endangered species persists. Kodilam and Luss [10] studies the allocation problem while minimizing a separable convex objective. Hochbaum[9] studies this a nonlinear integer or continuous knapsack maximization problem with a separable concave objective and a number of variants, providing efficient approximation algorithms for these problems. These are perhaps the most clear complexity results for this problem class. Bitran and Hax [1] study the resource allocation problem with a separable convex objective and suggest a recursive procedure to solve these problems. Bretthaur and Shetty[2, 3] study generalization of the nonlinear resource allocation problem with nonlinear constraints. They present an algorithm based on branch and bound. For more results and applications, see [11] for a survey of the continuous nonlinear resource allocation problem.

Contributions and outline Our main contribution (Theorem 13) is to show that Model 1 can be solved in constant number of operations, provided some numerical preprocessing (see Proposition 1) is done on the function f. We prove these results using only the basic assumptions discussed after Model 1. We prove a number of structural results along the way, which also imply a simple polynomial time algorithm (Corollary 4) to solve Model 1.

We begin with a KKT analysis of the optimal solutions which leads to the simple linear time algorithm. We then delve deeper into the structure of optimal solutions through a rephrasing of the problem in to a MINLP. Analysis of this problem reveals a finite number of potential optimal solutions to consider, which can be computed given proper preprocessing of the function. In particular, we need to compute a domain value d_a with specific properties related to f.

[1]Although we do not use it in this article, data related to these redistricting problems can be provided if requested.

2 Structure of KKT Solutions

Unless stated otherwise, we assume that $M \in [an, bn]$. Note that Model 1 is feasible if and only if $M \in [an, bn]$. Given a feasible solution $\mathbf{x}$ to Model 1, partition $I = \{1, \ldots, n\}$ into

$$I_{\mathbf{x}}^a = \{i \in I \ : \ x_i = a\}, \qquad I_{\mathbf{x}}^b = \{i \in I \ : \ x_i = b\}, \quad \text{and} \quad I_{\mathbf{x}}^y = \{i \in I \ : \ x_i \in (a,b)\}.$$

Lemma 2. *Let $\mathbf{x}^*$ be an optimal solution to Model 1. There exists a constant λ such that*

(a) $f'(x_i^*) = \lambda$ for every i in $I_{\mathbf{x}}^y$,

(b) If $I_{\mathbf{x}^*}^a$ is nonempty, then $f'(a) \leq \lambda$, and

(c) If $I_{\mathbf{x}^*}^b$ is nonempty, then $\lambda \leq f'(b)$.

Proof. The Lagrangian of Model 1 is given by

$$L(\mathbf{x}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b, \lambda) = \sum_{i=1}^n f(x_i) + \sum_{i=1}^n \mu_i^a (a - x_i) + \sum_{i=1}^n \mu_i^b (x_i - b) + \lambda \left(\sum_{i=1}^n M - x_i \right).$$

Notice that

$$\tfrac{\partial}{\partial x_i} \mathcal{L}(\mathbf{x}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b, \lambda) = f'(x_i) - \mu_i^a + \mu_i^b - \lambda.$$

Since $\mathbf{x}^*$ is optimal, the KKT necessary conditions tell us that, for each $i \in \{1, ..., n\}$,

$$\begin{cases} f'(x_i) = \lambda - \mu_i^a + \mu_i^b & \text{(3a)} \\ \mu_i^a (a - x_i^*) = 0 & \text{(3b)} \\ \mu_i^b (x_i^* - b) = 0 & \text{(3c)} \\ \boldsymbol{\mu}^a, \boldsymbol{\mu}^b \geq \mathbf{0}, & \text{(3d)} \end{cases}$$

If $i \in I_{\mathbf{x}^*}^y \cup I_{\mathbf{x}^*}^a$, then $a - x_i^* > 0$ so (3b) implies that $\mu_i^a = 0$. Similarly, if $i \in I_{\mathbf{x}^*}^y \cup I_{\mathbf{x}^*}^b$, then $x_i - b^* > 0$ so (3c) implies that $\mu_i^b = 0$. These conditions reduce to

$$\begin{cases} f'(x_i^*) = \lambda, & \forall\, i \in I_{\mathbf{x}^*}^y & \text{(4a)} \\ f'(a) = \lambda - \mu_i^a, & \forall\, i \in I_{\mathbf{x}^*}^a & \text{(4b)} \\ f'(b) = \lambda + \mu_i^b, & \forall\, i \in I_{\mathbf{x}^*}^b & \text{(4c)} \\ \boldsymbol{\mu}^a, \boldsymbol{\mu}^b \geq \mathbf{0}. & & \text{(4d)} \end{cases}$$

Notice that (4a) is exactly condition (a). Since $\mu_i^a \geq 0$, (4b) can only be satisfied if $f'(a) \leq \lambda$; similarly, (4c) can only be satisfied if $f'(b) \geq \lambda$. Thus, we have (b) and (c). Note that (4a), (4b), and (4c) only apply if their respective sets are nonempty. □

We call a solution $\mathbf{x}$ *trine* if each entry x_i takes one of three possible values. In particular, we use trine to describe solutions for which $x_i \in \{a, y, b\}$ for every i and some $y \in (a, b)$. That is, $x_i = y$ for every $i \in I_{\mathbf{x}}^y$. Additionally, we define the antisymmetric complement $\overline{x} = 2c - x$ of a point $x \in [a, b]$ so that $f'(x) = f'(\overline{x})$ by the antisymmetry of f.

Lemma 3. *Suppose (A1), (A2), (A3) hold. Then there is an optimal trine solution $\mathbf{x}^*$ to Model 1.*

Proof. It may be that c is not the midpoint of a and b; we will focus on the case where $b-c > c-a$ as the other cases follow by a similar argument.

In this case $\overline{a} < b$. By the antisymetry of f, we have $f(x) = 2f(c) - f(\overline{x})$ and thus

$$f(x) + f(\overline{x}) \quad = \quad 2f(c) - f(\overline{x}) + f(\overline{x}) \quad = \quad 2f(c)$$

for any $x \in [c, b]$.

Let $\mathbf{x}^*$ be an optimal, non-trine solution to Model 1; naturally, it must satisfy Lemma 2. By Lemma 2.a it must be that $f'(x_i^*)$ takes the same value for each $i \in I^y_{\mathbf{x}^*}$; call this value λ. Notice that since f' is strictly decreasing over $(c, b]$ there is at most one point y in $[c, b)$ which satisfies $f'(y) = \lambda$. However, by antisymmetry, $f'(\overline{y}) = f'(y) = \lambda$. Since $y \in [c, b)$, we know that $\overline{y} \le c$.

Since $\mathbf{x}^*$ is not trine but does satisfy Lemma 2, there exists a pair of distinct indices j, k such that $x_j^* = y$ and $x_k^* = \overline{y}$. Consider the perturbed solution $\mathbf{x}'$ defined by

$$x'_i = \begin{cases} y + \Delta, & \text{if } i = j \\ \overline{y} - \Delta, & \text{if } i = k \\ x_i, & \text{otherwise} \end{cases}$$

for some $\Delta \in [\overline{y} - c, \overline{y} - a] \setminus \{0\}$. Notice that $\overline{y} - \Delta = 2c - (y + \Delta) = \overline{y + \Delta}$ and thus

$$F(\mathbf{x}) - F(\mathbf{x}') \quad = \quad f(y) + f(\overline{y}) - f(y + \Delta) - f(\overline{y} - \Delta) \quad = \quad 2f(c) - 2f(c) \quad = \quad 0.$$

That is, we may perturb antisymmetric pairs in a solution without changing the objective function value. Further, such a permutation does not change $\sum_{i=1}^n x_i$ so the constraint $\sum_{i=1}^n x_i = M$ will not be violated and the bounds on Δ are chosen to prevent violation of $\mathbf{x} \in [a, b]^n$.

If $x_i \in \{a, b\}$ for all $i \neq j, k$, then such a perturbation does not change the satisfaction of Lemma 2.a and $\mathbf{x}'$ is also optimal. Letting $\Delta = \overline{y} - a$ gives $x'_j = 2c - a = \overline{a}$ and $x'_k = a$ so, in this case, $x'_i \in \{a, \overline{a}, b\}$ for each $i \in \{1, \dots, n\}$ so $\mathbf{x}'$ is trine. That is, we may perturb any non-trine, optimal solution with $|I^y_{\mathbf{x}^*}| = 2$ into a trine, optimal solution.

On the other hand, if there exists any $i \neq j, k$ for which $x_i \in \{y, \overline{y}\}$ then applying the perturbation for any Δ means that the entries of $\mathbf{x}'$ span either $\{a, \overline{y}, \overline{y} - \Delta, y + \Delta, b\}$ or $\{a, \overline{y} - \Delta, y + \Delta, y, b\}$. In either case, Lemma 2.a is violated and $\mathbf{x}'$ is not optimal, but $f(\mathbf{x}') = f(\mathbf{x})$ which contradicts our assumption that $\mathbf{x}$ is optimal. That is, no non-trine solution with $|I^y_{\mathbf{x}^*}| \ge 3$ is optimal. □

Define $\mathcal{X}(k_a, k_b)$ to be the family of trine solutions $\mathbf{x}$ to Model 1 such that $|I^a_{\mathbf{x}}| = k_a$ and $|I^b_{\mathbf{x}}| = k_b$. That is, k_a elements of $\mathbf{x}$ take the value a and k_b elements take the value b.

Define $k_y = n - k_a - k_b$ so that $|I^y_{\mathbf{x}}| = k_y$. Since each $\mathbf{x} \in \mathcal{X}(k_a, k_b)$ is trine, the constraint $\sum_{i=1}^n x_i = M$ tells us that

$$ak_a + bk_b + yk_y = M \quad \Rightarrow \quad y = \tfrac{M - ak_a - bk_b}{n - k_a - k_b}. \tag{5}$$

Thus each family defines a unique k_y and, assuming $k_y > 0$, a unique y.

Corollary 4 (Enumeration algorithm)**.** *There exists an algorithm that runs in $O(n^2)$ time and evaluations of f that computes the optimal solution to Model 1.*

Proof. Enumerate partitions of n into three integers, k_a, k_b, and k_y, of which there are $O(n^2)$ many such choices. Compute y via (5) (when $k_y \neq 0$). Then compare all values of the function $f(a)k_a + f(b)k_b + f(y)k_y$ and return the largest. □

Note that this algorithm can be improved to $O(n)$ time complexity when you can argue that either k_a or k_b is zero in an optimal solution.

We will use this k-space perspective in the following section.

3 A k-space model

We can express the objective function value of a solution in terms of our new parameters:

$$F(\mathbf{x}) = f(a)k_a + f(b)k_b + f(y)k_y,$$

for any $\mathbf{x} \in \mathcal{X}(k_a, k_b)$. Ergo, solutions are homomorphic within their family and we may define a new but reduced model which exploits this structure.

Model 5 (k-space model)**.**

$$\begin{aligned} \text{Maximize} \quad & \mathcal{F}(\mathbf{k}) = f(a)k_a + f(b)k_b + f(y)k_y && \text{(6a)} \\ \text{s.t.} \quad & ak_a + bk_b + yk_y = M && \text{(6b)} \\ & k_a + k_b + k_y = n && \text{(6c)} \\ & k_a, k_b, k_y \in \mathbb{Z}_+ && \text{(6d)} \\ & y \in (a, b) && \text{(6e)} \end{aligned}$$

A solution to Model 5 takes the form (k_a, k_b, k_y, y), is uniquely defined by any two of its elements, and represents a family of trine solutions to Model 1 (if $k_y = 0$, y does not contribute to the satisfaction of (6b) or to the objective function and so can take any value). We proceed by solving the continuous relaxation of this MINLP.

3.1 Continuous Relaxation

We call a solution $\mathbf{k} = (k_a, k_b, k_y, y)$ *equality-feasible* if it satisfies the equality constraints (6b) and (6c) of Model 5, but not necessarily non-negativity, integrality (6d), or the bounds on y (6e).

Lemma 6 (Feasibility in terms of k_a and k_b)**.** *An equality feasible solution (k_a, k_b, k_y, y) is fully feasible to the continuous relaxation of Model 5 if and only if*

$$\frac{yn - M}{y - a} \leq k_a \leq \frac{bn - M}{b - a} \quad \textit{and} \quad \frac{M - yn}{b - y} \leq k_b \leq \frac{M - an}{b - a}.$$

Proof. Constraints (6c) and (6d) imply

$$0 \leq k_a, k_b, k_y \leq n \tag{7}$$

Since this solution is equality-feasible, we may solve (6b) and (6c) for k_a and k_y:

$$\begin{aligned} k_a + k_b + k_y = n \quad & \Rightarrow \quad k_a = n - k_b - k_y, && \text{(8a)} \\ ak_a + bk_b + yk_y = M \quad & \Rightarrow \quad k_y = \frac{(M - an) - (b - a)k_b}{y - a}. && \text{(8b)} \end{aligned}$$

Applying these equations to the implied constraint (7) gives

$$\begin{array}{ccccccccc} 0 & \leq & k_a & \leq & n & \Rightarrow & 0 \leq \frac{(M-yn)+(b-y)k_b}{y-a} \leq n \\ & & & & & & \frac{M-yn}{b-y} \leq k_b \leq \frac{M-an}{b-y} \\ 0 & \leq & k_y & \leq & n & \Rightarrow & 0 \leq \frac{(M-an)-(b-a)k_b}{y-a} \leq n \\ & & & & & & \frac{M-an}{b-a} \geq k_b \geq \frac{M-yn}{b-a} \end{array}$$

but $b-a > b-y$, so

$$\tfrac{M-yn}{b-a} \quad < \quad \tfrac{M-yn}{b-y} \quad \leq \quad k_b \quad \leq \quad \tfrac{M-an}{b-a} \quad < \quad \tfrac{M-an}{b-y}.$$

Solving constraints (6b) and (6c) for k_b and k_y and performing similar operations gives

$$\tfrac{yn-M}{b-a} \quad < \quad \tfrac{yn-M}{y-a} \quad \leq \quad k_a \quad \leq \quad \tfrac{bn-M}{b-a} \quad < \quad \tfrac{bn-M}{y-a}.$$

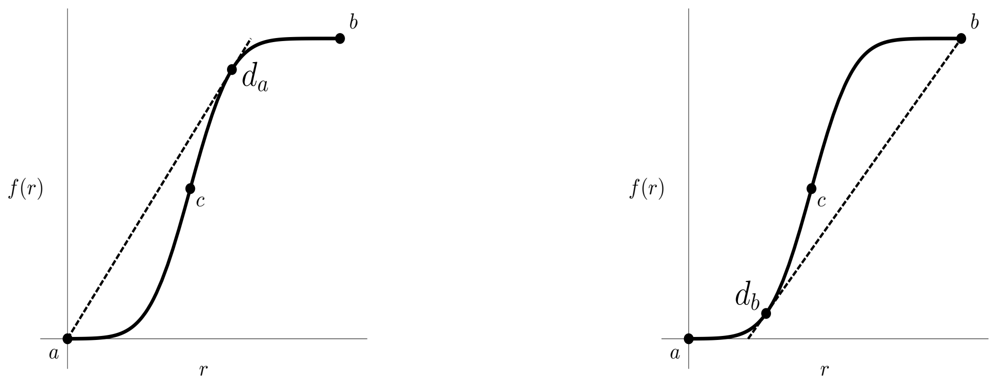
Further, both upper bounds are corroborated applying (5) to (6e). □

Lemma 7. *For any $r \in [a,b] \setminus \{c\}$, the equation*

$$f(r) = f(x) + (r-x)f'(x) \tag{9}$$

has exactly two solutions: $x = r$ and another, which lies between $\overline{r} = 2c - r$ and c. We denote this additional solution d_r.

Of particular interest are the values d_a and d_b. See Figure 2 for a graphical definition of these two points.



(a) Tangent definition of d_a (b) Tangent definition of d_b

Figure 2: d_r is the point at which the tangent line of f passes through $(r, f(r))$

Proof. Because f is antisymmetric about c, it suffices to show these properties for $r < c$.

Define $g(x) := f(x) + (r-x)f'(x) - f(r)$. Clearly $g(r) = 0$, so $x = r$ is a solution to the equation.

We claim that $g(c) < 0$ and $g(2c - r) > 0$. Thus, by the Intermediate Value Theorem, we conclude that there exists a point $d_r \in (2c - r, c)$ such that $g(d_r) = 0$.

First note that since $f(x)$ is strictly convex over $[a, c]$, the first-order-convexity of f at $x = c$ implies

$$g(c) = f(x) + (r-x)f'(r) - f(r) < 0$$

for any $x \in [c, b]$. In particular $g(c) < 0$.

Now notice that

$$\begin{aligned} g(2c-r) &= f(2c-r) + (r-(2c-r))f'(2c-r) - f(r) \\ &= 2f(c) + 2(r-c)f'(r) - 2f(r) && \text{by assumption (A2)} \\ &= -2\left(f(r) + (c-r)f'(r) - f(c)\right) > 0, \end{aligned}$$

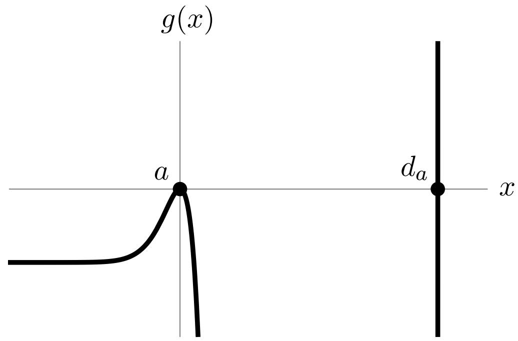


Figure 3: A plot of $g(x)$ in the proof Lemma 7 for $r = a$. In particular, an analysis of $g(x)$ and $g'(x)$ show that there are exactly two roots which are at $x = a$ and $x = d_a$.

where the last inequality follows from first-order-convexity of f applied at $x = r$.

$$g'(x) = (r - x)f''(x)$$

which since f is strictly convex over $[a, c]$ and strictly concave over $[c, b]$

$$\begin{cases} g'(x) < 0 & \text{if } x \in [a, r), \\ g'(x) > 0 & \text{if } x \in (r, c), \\ g'(c) = 0, & \\ g'(x) < 0 & \text{if } x \in (c, b], \end{cases}$$

which implies that our roots are unique. See Figure 3. □

Proposition 1. For any $r \in [a, b]$, we can compute d_r up to ϵ additive error in $O(\log(\frac{1}{\epsilon}))$ many oracle calls of f and f'.

Proof. Define $g(x)$ as in the previous proof. By the previous lemma, $g(r) = 0$ and $g(d_r) = 0$. Furthermore, d_r lies between $2c - r$ and c. Since d_r is unique, g doesn't have any other zeros. Thus, we can run the bisection method with initial bounds $2c - r$ and c to obtain the desired complexity. Note that each evaluation of g requires evaluations of both f' and f. □

Theorem 8 (KKT Solutions to Model 5)**.** *The KKT solutions to the Continuous Relaxation of Model 5 are given in the following table.*

Case	***Condition***	k_a	k_b	k_y	y
0	$k_y = 0$	$\frac{bn-M}{b-a}$	$\frac{M-an}{b-a}$	0	n/a
1	$k_a, k_b, k_y > 0$	*No KKT Solution*			
2A	$k_y = n$	0	0	n	$\frac{M}{n}$
2B	$k_b = 0$	$\frac{d_a n-M}{d_a-a}$	0	$\frac{M-an}{d_a-a}$	d_a
2C	$k_a = 0$	0	$\frac{M-d_b n}{b-d_b}$	$\frac{bn-M}{b-d_b}$	d_b

Proof. The Lagrangian function is

$$\begin{aligned} \mathcal{L}(\mathbf{k}, \lambda, \mu) = f(a)k_a + f(b)k_b + f(y)k_y + \lambda_1(ak_a + bk_b + yk_y - M) + \lambda_2(k_a + k_b + k_y - n) \\ + \mu_a(k_a - 0) + \mu_b(k_b - 0) + \mu_y(k_y - 0) + \xi^-(y - a) + \xi^+(b - y) \end{aligned} \tag{10}$$

The corresponding KKT conditions are

$$\begin{aligned}
f(a) + \lambda_1 a + \lambda_2 + \mu_a &= 0 && \text{(11a)}\\
f(b) + \lambda_1 b + \lambda_2 + \mu_b &= 0 && \text{(11b)}\\
f(y) + \lambda_1 y + \lambda_2 + \mu_y &= 0 && \text{(11c)}\\
k_y f'(y) + \lambda_1 k_y + \xi^- - \xi^+ &= 0 && \text{(11d)}\\
\forall\, i \in \{a,b,y\} \quad \mu_i \cdot k_i &= 0 && \text{(11e)}\\
\xi^-(y-a) &= 0 && \text{(11f)}\\
\xi^+(b-y) &= 0 && \text{(11g)}
\end{aligned}
\qquad
\begin{aligned}
ak_a + bk_b + yk_y &= M && \text{(11h)}\\
k_a + k_b + k_y &= n && \text{(11i)}\\
k_a, k_b, k_y &\geq 0 && \text{(11j)}\\
a \leq y &\leq b && \text{(11k)}\\
\forall\, i \in \{a,b,y\} \quad \mu_i &\geq 0 && \text{(11l)}\\
\xi^-, \xi^+ &\geq 0 && \text{(11m)}
\end{aligned}$$

0A. ($k_y = 0$) Then, by (11h) and (11i), the system must satisfy these equations:

$$k_a + k_b = n$$
$$ak_a + bk_b = M.$$

By assumption, $a \neq b$, and hence the solution for $k_a = \frac{bn-M}{b-a}$ and $k_b = \frac{M-an}{b-a}$ is unique.

0B. ($y = a$ or $y = b$) Then k_y can be combined with either k_a or k_b, again reducing the problem to these two variables. Hence, this has the same outcome as Case **0A**.

We may now assume that $a < y < b$ and $k_y > 0$. Hence, $\xi^-, \xi^+ = 0$ by (11g) and (11f). By (11d),

$$k_y f'(y) + \lambda_1 k_y = 0 \quad \Rightarrow \quad \lambda_1 = -f'(y)$$

Then by (11a), (11b), and (11c), solving for λ_2, we have

$$f(a) - f'(y)a + \mu_a \quad = \quad f(b) - f'(y)b + \mu_b \quad = \quad f(y) - f'(y)y + \mu_y$$

which can be rewritten as

$$f(b) - f(a) - f'(y)(b-a) + (\mu_b - \mu_a) = 0, \tag{12a}$$
$$f(b) - f(y) - f'(y)(b-y) + (\mu_b - \mu_y) = 0, \tag{12b}$$
$$f(y) - f(a) - f'(y)(y-a) + (\mu_y - \mu_a) = 0. \tag{12c}$$

Note that one of these equations is redundant.

1. ($k_a, k_y, k_b > 0$) Then by (11e), $\mu_i = 0$ for $i \in \{a,b,y\}$ and equations (12b) and (12c) give

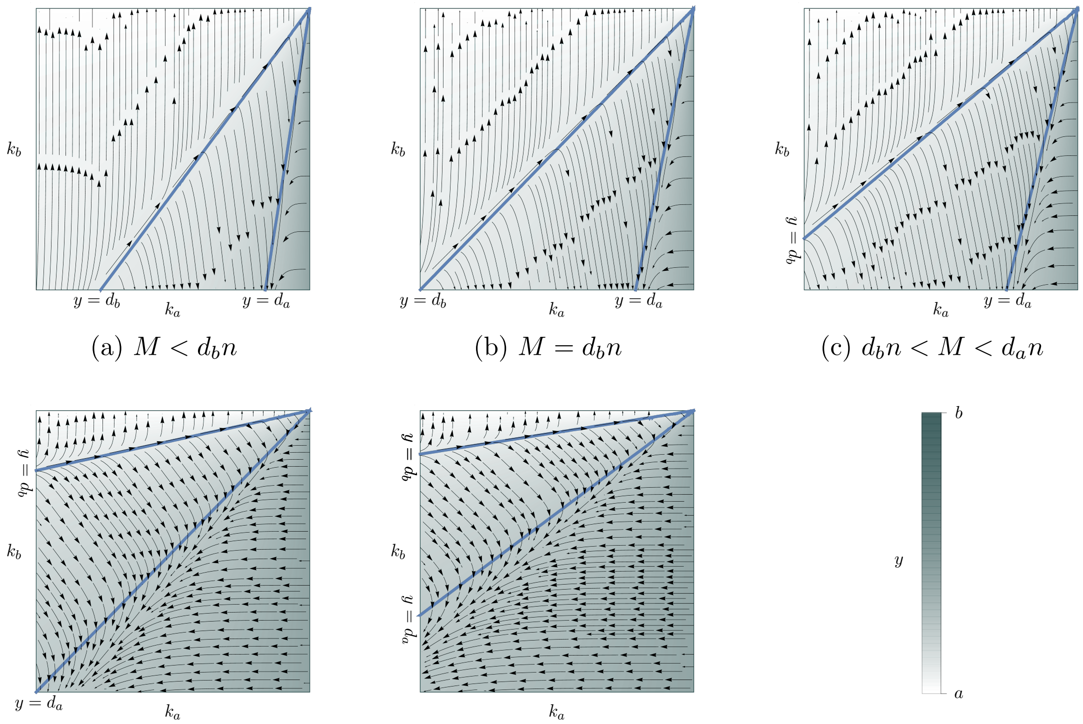
$$f(b) - f(y) - f'(y)(b-y) = 0 \qquad \text{and} \qquad f(y) - f(a) - f'(y)(y-a) = 0.$$

According to Lemma 7, the unique solutions to these equations, respectively, are $y = d_b$ and $y = d_a$ ($y = a$ and $y = b$ are also roots but have already been excluded by case **0B**). Also by Lemma 7, it must be that $d_b < c < d_a$ because $a < c < b$. Thus there is no KKT appropriate solution in this case.

2A. ($k_y = n$) Then simply $y = \frac{M}{n}$, $k_a = k_b = 0$.

2B. ($k_b = 0$, $k_a > 0$, $k_y > 0$) Then $\mu_1 = \mu_3 = 0$. Hence, by (12c),

$$f(y) - f(a) - f'(y)(y-a) = 0.$$



(a) $M < d_b n$ (b) $M = d_b n$ (c) $d_b n < M < d_a n$

(d) $M = d_a n$ (e) $M > d_a n$ (f) Coloring scale for y

Figure 4: Stream Plots of the Gradient Field of $F(k_a, k_b)$. Various cases are exhibited that expose different KKT solutions as possible optimal solutions.

By Lemma 9, there is a unique solution, $y = d_a$. Hence

$$ak_a + d_a k_y = M \qquad \text{and} \qquad k_a + k_y = n.$$

Solving this system of equations yields

$$y = d_a, \qquad k_y = \frac{M - an}{d_a - a}, \qquad \text{and} \qquad k_a = n - \frac{M - an}{d_a - a} = \frac{d_a n - M}{d_a - a}.$$

2C. ($k_a = 0$, $k_b > 0$, $k_y > 0$) Similar to Case **2B**, we have the solution

$$y = d_b, \qquad k_y = \frac{bn - M}{b - d_b}, \qquad \text{and} \qquad k_b = n - \frac{bn - M}{b - d_b} = \frac{M - d_b n}{b - d_b}.$$

This concludes the case analysis. □

Using (5), we can uniquely project an equality-feasible solution $\mathbf{k} = (k_a, k_b, k_y, y)$ onto any two of its parameters. Let $\mathcal{F}(k_a, k_b)$ be the projection of $F(\mathbf{k})$ onto the (k_a, k_b) space. Under this projection, both y and k_y are functions of k_a and k_b. Figure 4 contains plots the vector field of the $\nabla\mathcal{F}(k_a, k_b)$ (the value of y is represented by background color gradient). Theorems 9 and 10 are informed by Figure 4.

Theorem 9. *Let $\mathbf{k} = (k_a, k_b, k_y, y)$ be an equality-feasible solution. We have*

$$\frac{\partial}{\partial k_a} y \geq 0 \qquad \text{and} \qquad \begin{cases} \frac{\partial}{\partial k_a}\mathcal{F}(k_a, k_b) = 0, & \text{if } y = d_a, \\ \frac{\partial}{\partial k_a}\mathcal{F}(k_a, k_b) > 0, & \text{if } y \in (a, d_a), \\ \frac{\partial}{\partial k_a}\mathcal{F}(k_a, k_b) < 0, & \text{if } y \in (d_a, b). \end{cases}$$

Proof. Recall from equation (5) that since $\mathbf{k}$ is equality-feasible, we can treat y and k_y as functions of k_a and k_b from (6b) and (6c)

$$k_y = n - k_a - k_b \qquad \text{and} \qquad y = \frac{M - ak_a - bk_b}{k_y}.$$

Clearly $\frac{\partial}{\partial k_a} k_y = \frac{\partial}{\partial k_a} n - k_a - k_b = -1$. It follows that

$$\frac{\partial}{\partial k_a} y \quad = \quad \frac{\partial}{\partial k_a}\left(\frac{M - ak_a - bk_b}{k_y}\right) \quad = \quad \frac{-a}{k_y} + \frac{M - ak_a - bk_b}{k_y^2} \quad = \quad \frac{y-a}{k_y}.$$

which is non-negative for any $y \in (a, b)$ and $k_y > 0$.

Now, notice that the partial derivative

$$\begin{aligned}\frac{\partial}{\partial k_a}\mathcal{F}(k_a, k_b) &= \frac{\partial}{\partial k_a}\left(f(a)k_a + f(b)k_b + f(y)k_y\right) = f(a) + f'(y)\left(\frac{\partial}{\partial k_a} y\right) k_y + f(y)\frac{\partial}{\partial k_a} k_y \\ &= f(a) + f'(y)\,(y-a) - f(y)\end{aligned}$$

equals zero if $y = a$ or $y = d_a$. These are the only roots since d_a is unique.

Suppose that $y = c$. Since $f(x)$ is strictly convex for $x \in [a, c)$, we have $f(a) > f(c) - f'(c)(c-a)$, so $\frac{\partial}{\partial k_a}\mathcal{F}(k_a, k_b)$ is positive. Now suppose that $y = \overline{a}$. By the antisymmetry of f, we have

$$f(\overline{a}) \quad = \quad 2f(c) - f(a) \quad < \quad 2f(c) - f(c) + f'(c)(c-a) \quad = \quad f(c) - f'(c)(c - \overline{a}),$$

so $\frac{\partial}{\partial k_a}\mathcal{F}(k_a, k_b) < 0$. Since $d_a \in (c, \overline{a})$ and there are no other roots in (a, b), the result holds. □

This theorem implies that: given a solution with $y < d_a$, increasing k_a increases both y and $\mathcal{F}(\mathbf{k})$; or given a solution $y > d_a$, increasing k_a decreases y but increases $\mathcal{F}(\mathbf{k})$. In either case, we can perturb k_a to any feasible extent because y will not leave the region in question. We can see this behavior from Figure 4 by noticing that the k_a component of each vector points towards the line $y = d_a$. Similarly, the k_b component of each vector points away from the line $y = d_b$ which informs the next theorem.

Theorem 10. *Let $\mathbf{k} = (k_a, k_b, k_y, y)$ be an equality-feasible solution. We have*

$$\frac{\partial}{\partial k_b} y \le 0 \qquad \text{and} \qquad \begin{cases} \frac{\partial}{\partial k_b}\mathcal{F}(\mathbf{k}) = 0, & \text{if } y = d_b \\ \frac{\partial}{\partial k_b}\mathcal{F}(\mathbf{k}) > 0, & \text{if } y \in (a, d_b) \\ \frac{\partial}{\partial k_b}\mathcal{F}(\mathbf{k}) < 0, & \text{if } y \in (d_b, b). \end{cases}$$

The proof of this parallels that of Theorem 9. With the behavior of the gradient now well understood, we can find the optimal solution.

Corollary 11 (Ordering and feasibility of KKT solutions)**.** *The optimal solution to Model 5 is given by*

$$(k_a^*, k_b^*, k_y^*, y^*) = \begin{cases} \left(\frac{d_a n - M}{d_a - a}, 0, \frac{M - an}{d_a - a}, d_a\right) & \text{if } M < d_a n, \\ \left(0, 0, n, \frac{M}{n}\right) & \text{if } M \ge d_a n. \end{cases}$$

Proof. First note that, $\mathbf{k^{2A}}$ is always feasible, while $\mathbf{k^{2B}}$ is only feasible when $M \le d_a n$. Hence, it suffices to prove the ordering only considering equality-feasibility.

$\left[\mathcal{F}\left(\mathbf{k^{2B}}\right) \geq \mathcal{F}\left(\mathbf{k^{2A}}\right)\right]$ By Theorem 9,

$$\mathcal{F}\left(\mathbf{k^{2B}}\right) = \mathcal{F}\left(\frac{d_a n - M}{d_a - a}, 0, \frac{M - an}{d_a - a}, d_a\right) \geq \mathcal{F}\left(0, 0, n, \frac{M}{n}\right) = \mathcal{F}\left(\mathbf{k^{2A}}\right)$$

since k_b is kept constant. (*Note:* these solutions are equivalent if and only if $M = d_a n$.)

$\left[\mathcal{F}\left(\mathbf{k^{2B}}\right) > \mathcal{F}\left(\mathbf{k^{0}}\right)\right]$ By Theorem 9

$$\mathcal{F}\left(\mathbf{k^{2B}}\right) = \mathcal{F}\left(\frac{d_a n - M}{d_a - a}, 0, \frac{M - an}{d_a - a}, d_a\right) > \mathcal{F}\left(\frac{bn - M}{b - a}, 0, \frac{M - an}{b - a}, y^*\right)$$

since k_b is kept constant. By (5), $y^* = b$ so we can exchange k_y for k_b and the latter solution is actually $\mathbf{k^0}$.

$\left[\mathcal{F}\left(\mathbf{k^{0}}\right) > \mathcal{F}\left(\mathbf{k^{2C}}\right)\right]$ By Theorem 10,

$$\mathcal{F}\left(\mathbf{k^{2C}}\right) = \mathcal{F}\left(0, \frac{M - d_b n}{b - d_b}, \frac{bn - M}{b - d_b}, d_b\right) < \mathcal{F}\left(0, \frac{M - an}{b - a}, \frac{bn - M}{b - a}, y^*\right)$$

since k_a is kept constant. By (5), $y^* = a$ and the latter solution is actually $\mathbf{k^0}$.

$\left[\textbf{If } M \geq d_a n\textbf{, then } \mathcal{F}\left(\mathbf{k^{2A}}\right) > \mathcal{F}\left(\mathbf{k}^{0}\right)\textbf{.}\right]$ By Theorem 9,

$$\mathcal{F}\left(\mathbf{k^{2A}}\right) = \mathcal{F}\left(0, 0, n, \frac{M}{n}\right) > \mathcal{F}\left(\frac{bn - M}{b - a}, 0, \frac{M - an}{b - a}, y^*\right)$$

since $\frac{M}{n} \geq d_a$ and k_b is kept constant. By (5), $y^* = b$ and the latter solution is actually $\mathbf{k^{2B}}$. □

3.2 Integer Solutions

Corollary 12 (to Lemma 6)**.** *A equality-feasible solution* (k_a, k_b, k_y, y) *is feasible to Model 5 if and only if*

$$\left\lceil \tfrac{yn-M}{y-a} \right\rceil \leq k_a \leq \left\lfloor \tfrac{bn-M}{b-a} \right\rfloor \qquad \textit{and} \qquad \left\lceil \tfrac{M-yn}{b-y} \right\rceil \leq k_b \leq \left\lfloor \tfrac{M-an}{b-a} \right\rfloor$$

We define the following points for the next theorem:

Theorem 13. *The optimal integer solution to Model 5 is given by*

$$(k_a^*, k_b^*, k_y^*, y^*) = \begin{cases} \textit{best of } 0_-,\ \boldsymbol{2B}_-,\ \boldsymbol{2B}^+ & \textit{if } M < d_a n, \\ \textit{Case } \boldsymbol{2A} & \textit{if } M \geq d_a n. \end{cases} \tag{13}$$

Thus, an optimal solution can be computed with a constant number of operations, provided a pre-processing algorithm to determine d_a *with sufficient accuracy.*

Proof. According to Corollary 11, solution **2A** is the optimal solution to the continuous relaxation if $M \geq d_a n$. Since it is integral, it must also be optimal to the integer problem. Thus, we now assume that $M < d_a n$ and aim to show that any integer feasible solution $\mathbf{k} = (k_a, k_b, k_y, y)$ is dominated by at least one of $\mathbf{k}^{0-}$, $\mathbf{k^{2B}}_-$, and $\mathbf{k^{2B}}^+$. We start by noticing that there are two cases, $y \leq d_b$ and $y \geq d_b$.

Point	k_a	k_b	k_y	y
$\mathbf{0}_-$	$\left\lfloor \frac{bn-M}{b-a} \right\rfloor$	$\left\lfloor \frac{M-an}{b-a} \right\rfloor$	$n-k_a-k_b$	$\star\, d_a$
2A	0	0	n	$\frac{M}{n}$
$\mathbf{2B}_-$	$\left\lfloor \frac{d_a n-M}{d_a-a} \right\rfloor$	0	$\left\lceil \frac{M-an}{d_a-a} \right\rceil$	$< d_a$
$\mathbf{2B}^+$	$\left\lceil \frac{d_a n-M}{d_a-a} \right\rceil$	0	$\left\lfloor \frac{M-an}{d_a-a} \right\rfloor$	$> d_a$

Table 1: Four potentially optimal integer solutions, where the columns explain the values of the variables from the Model 5. The labels in the left column correspond to case analysis in proof of Theorem 8. Having a finite list of possible solutions allows us to compute the optimal integer solution via a constant number of operations. See Theorem 13.

(I) ($y \leq d_b$) By Theorem 10, $\mathcal{F}(k_a, k_b, k_y, y) < \mathcal{F}\left(k_a, \left\lfloor \frac{M-an}{b-a} \right\rfloor, k_y^*, y^*\right)$ since k_a is left constant. Notice that k_b has been increased to its maximum integer-feasible value, so $y* < y \leq d_b < d_a$. Thus Theorem 9 tells us that

$$\mathcal{F}\left(k_a, \left\lfloor \frac{M-an}{b-a} \right\rfloor, k_y^*, y^*\right) < \mathcal{F}\left(\left\lfloor \frac{bn-M}{b-a} \right\rfloor, \left\lfloor \frac{M-an}{b-a} \right\rfloor, k_y^\circ, y^\circ\right).$$

This last solution is exactly $\mathbf{k}^{0_-}$.

(II) ($y \geq d_b$) By Theorem 10,

$$\mathcal{F}(k_a, k_b, k_y, y) < \mathcal{F}\left(k_a, 0, k_y^*, y^*\right)$$

since k_a is left constant. Notice that k_b is decreased, which means that y has increased, so we could have either $y^* \geq d_a$ or $y^* \leq d_a$.

(i) ($y^* \geq d_a$) In this case, since $k_b = 0$ (6b) gives

$$M = ak_a + y(n-k_a) \geq ak_a + d_a(n-k_a)$$

which gives $k_a \geq \frac{d_a n-M}{d_a-a}$. However, because $k_a \in \mathbb{Z}$ it must be that $k_a \geq \left\lceil \frac{d_a n-M}{d_a-a} \right\rceil$. With this, Theorem 9 tells us that $\mathcal{F}\left(k_a, 0, k_y^*, y^*\right) \leq \mathcal{F}\left(\left\lceil \frac{d_a n-M}{d_a-a} \right\rceil, 0, k_y^\circ, y^\circ\right)$ which is exactly $\mathbf{k}^{\mathbf{2B}^+}$

(ii) ($y^* \leq d_a$) In this case, by similar logic as above, $k_a \leq \left\lfloor \frac{d_a n-M}{d_a-a} \right\rfloor$. With this, Theorem 9 tells us that $\mathcal{F}\left(k_a, 0, k_y^*, y^*\right) < \mathcal{F}\left(\left\lfloor \frac{d_a n-M}{d_a-a} \right\rfloor, 0, k_y^\circ, y^\circ\right)$.

Finally, comparing the objective value of these candidates, we can find the optimal solution. □

Given an optimal solution $(k_a^*, k_b^*, k_y^*, y^*)$ to Model 5, any $\mathbf{x} \in \mathcal{X}(k_a, k_b)$ is an optimal solution to Model 1.

Acknowledgements We would like to thank Nicholas Goedert for inspiration for this paper.

References

[1] Gabriel R. Bitran and Arnoldo C. Hax. Disaggregation and resource allocation using convex knapsack problems with bounded variables. *Management Science*, 27(4):431–441, 1981.

[2] Kurt M. Bretthauer and Bala Shetty. The nonlinear resource allocation problem. *Operations Research*, 43(4):670–683, 1995.

[3] Kurt M. Bretthauer and Bala Shetty. A pegging algorithm for the nonlinear resource allocation problem. *Computers Operations Research*, 29(5):505–527, 2002.

[4] Cyrus Derman. A simple allocation problem. *Management Science*, 5(4):453–459, 1959.

[5] Jamie Fravel, Robert Hildebrand, Nicholas Goedert, Laurel Travis, and Matthew Pierson. Dual bounds for redistricting problems with non-linear objective functions. 2022.

[6] Nicholas Goedert, Jamie Fravel, Robert Hildebrand, Laurel Travis, and Matthew Pierson. Black representation and district compactness in southern congressional districts. 2022.

[7] Nicholas Goedert, Jamie Fravel, Robert Hildebrand, Laurel Travis, and Matthew Pierson. Comparison of new congressional district maps to optimized algorithmic maps. 2022.

[8] Robert G. Haight and Laurel E. Travis. Reserve design to maximize species persistence. *Environmental Modeling Assessment*, 13:243–253, 2008.

[9] Dorit S. Hochbaum. A nonlinear knapsack problem. *Operations Research Letters*, 17(3):103–110, 1995.

[10] Muralidharan S. Kodialam and Hanan Luss. Algorithms for separable nonlinear resource allocation problems. *Operations Research*, 46(2):272–284, 1998.

[11] Michael Patriksson. A survey on the continuous nonlinear resource allocation problem. *European Journal of Operational Research*, 185(1):1–46, 2008.