STATISTICS IN TRANSITION new series, December 2023
Vol. 24, No. 5, pp. 203–219, https://doi.org/10.59170/stattrans-2023-071
Received – 22.03.2022; accepted – 16.06.2023

Another solution for some optimum allocation problem

Wojciech Wójciak[1]

Abstract

We derive optimality conditions for the optimum sample allocation problem in stratified sampling, formulated as the determination of the fixed strata sample sizes that minimize the total cost of the survey, under the assumed level of variance of the stratified π estimator of the population total (or mean) and one-sided upper bounds imposed on sample sizes in strata. In this context, we presume that the variance function is of some generic form that, in particular, covers the case of the simple random sampling without replacement design in strata. The optimality conditions mentioned above will be derived from the Karush-Kuhn-Tucker conditions. Based on the established optimality conditions, we provide a formal proof of the optimality of the existing procedure, termed here as *LRNA*, which solves the allocation problem considered. We formulate the *LRNA* in such a way that it also provides the solution to the classical optimum allocation problem (i.e. minimization of the estimator's variance under a fixed total cost) under one-sided lower bounds imposed on sample sizes in strata. In this context, the *LRNA* can be considered as a counterparty to the popular recursive Neyman allocation procedure that is used to solve the classical problem of an optimum sample allocation with added one-sided upper bounds. Ready-to-use R-implementation of the *LRNA* is available through our `stratallo` package, which is published on the Comprehensive R Archive Network (CRAN) package repository.

Key words: stratified sampling, optimum allocation, minimum cost allocation under upper bounds, optimum allocation constant variance, optimum allocation under lower bounds, recursive Neyman algorithm.

1. Introduction

Let us consider a finite population U consisting of N elements. Let the parameter of principal interest of a single study variable y in U be denoted by θ. This parameter is the population total (i.e. $\theta = \sum_{k\in U} y_k$, where y_k denotes the value of y for population element $k \in U$), or the population mean (i.e. $\theta = \frac{1}{N}\sum_{k\in U} y_k$). To estimate θ, we consider the *stratified π estimator*, i.e. the *π estimator* of Horvitz and Thompson (see, e.g. Särndal, Swensson and Wretman, 1992, Section 2.8, p. 42) in *stratified sampling*. Under this well-known sampling technique, population U is stratified, i.e. $U = \bigcup_{h\in\mathcal{H}} U_h$, where U_h, $h \in \mathcal{H}$, called strata, are pairwise disjoint and non-empty, and $\mathcal{H} = \{1, \ldots, H\}$ denotes a finite set of strata indices of size $H \geq 1$. The size of stratum U_h is denoted N_h, $h \in \mathcal{H}$ and clearly $\sum_{h\in\mathcal{H}} N_h = N$. Probability samples of size $n_h \leq N_h$, $h \in \mathcal{H}$ are selected independently from each stratum according to chosen sampling designs, which are often the same in all strata. The resulting total sample is of size $n = \sum_{h\in\mathcal{H}} n_h \leq N$. It is well know that the *stratified π estimator* $\hat{\theta}$ of θ and its variance $V_{\hat{\theta}}$ are expressed in terms of the first

[1]Warsaw University of Technology, Poland. E-mail: wojciech.wojciak.dokt@pw.edu.pl.
ORCID: https://orcid.org/0000-0002-5042-160X.

and second order inclusion probabilities (see, e.g. Särndal et al. (1992, Result 3.7.1, p. 102) for the case when θ is the population total). In particular, for several important sampling designs

$$V_{\hat{\theta}}(\mathbf{n}) = \sum_{h \in \mathcal{H}} \frac{A_h^2}{n_h} - A_0, \tag{1}$$

where $\mathbf{n} = (n_h,\, h \in \mathcal{H})$ and $A_0,\, A_h > 0,\, h \in \mathcal{H}$ do not depend on $\mathbf{n}$. Among the most basic and common examples that give rise to the variance of the form (1) is the *stratified* π *estimator* of the population total with *simple random sampling without replacement* design in strata. This case yields in (1): $A_h = N_h S_h,\, h \in \mathcal{H}$, and $A_0 = \sum_{h \in \mathcal{H}} N_h S_h^2$, where S_h denotes stratum standard deviation of study variable y (see, e.g. Särndal et al., 1992, equation 3.7.8, p. 103).

The values of the strata sample sizes $n_h,\, h \in \mathcal{H}$, are chosen by the sampler. They may be selected to minimize the variance (1) at the admissible level of the total cost of the survey or to minimize the total cost of the survey subject to a fixed precision (1). The simplest total cost function is of the form:

$$c(\mathbf{n}) = c_0 + \sum_{h \in \mathcal{H}} c_h n_h, \tag{2}$$

where c_0 is a fixed overhead cost and $c_h > 0$ is the cost of surveying one element in stratum $U_h,\, h \in \mathcal{H}$. For further references, see, e.g. Särndal et al. (1992, Section 3.7.3, p. 104) or Cochran (1977, Section 5.5, p. 96). In this paper, we are interested in the latter strategy, i.e. the determination of the sample allocation $\mathbf{n}$ that minimizes total cost (2) under assumed fixed level of the variance (1). We also impose one-sided upper bounds on sample sizes in strata. Such optimization problem can be conveniently written in the language of mathematical optimization as Problem 1.1, in the definition of which we intentionally omit fixed overhead cost c_0 as it has no impact on the optimal solution to this problem.

Problem 1.1. Given a finite set $\mathcal{H} \neq \emptyset$ and numbers $A_0,\, A_h > 0,\, c_h > 0,\, M_h > 0$, such that $M_h \leq N_h,\, h \in \mathcal{H}$, and $V \geq \sum_{h \in \mathcal{H}} \frac{A_h^2}{M_h} - A_0 \geq 0$,

$$\underset{\mathbf{x} = (x_h,\, h \in \mathcal{H}) \in \mathbb{R}_+^{|\mathcal{H}|}}{\text{minimize}} \quad \sum_{h \in \mathcal{H}} c_h x_h \tag{3}$$

$$\text{subject to} \quad \sum_{h \in \mathcal{H}} \frac{A_h^2}{x_h} - A_0 = V \tag{4}$$

$$x_h \leq M_h, \qquad h \in \mathcal{H}. \tag{5}$$

To emphasize the fact that the optimal solution to Problem 1.1 may not be an integer one, we denote the optimization variable by $\mathbf{x}$, not by $\mathbf{n}$. Non-integer solution can be rounded up in practice with the resulting variance (1) being possibly near V, instead of the exact V. The upper bounds M_h imposed on $x_h,\, h \in \mathcal{H}$, are natural since for instance the allocation with $x_h > N_h$ for some $h \in \mathcal{H}$ is impossible. We assume that $V \geq \sum_{h \in \mathcal{H}} \frac{A_h^2}{M_h} - A_0$, since otherwise, if $V < \sum_{h \in \mathcal{H}} \frac{A_h^2}{M_h} - A_0$, the problem is infeasible. We also note that in the case when $V = \sum_{h \in \mathcal{H}} \frac{A_h^2}{M_h} - A_0$, the solution is trivial, i.e.: $\mathbf{x}^* = (M_h,\, h \in \mathcal{H})$.

It is worth noting that in the definition of Problem 1.1, we require (through (4)) that the variance defined in (1) is equal to a certain fixed value, denoted as V, and not less than that, whilst, it might seem more favourable at first, to require the variance (1) to be less than or equal to V, especially given the practical context in which Problem 1.1 arises (see Section 2 below). It is easy to see, however, that the objective function (3) and the variance constraint (4) are of such a form that the minimum of (3) is achieved for a value that yields the allowable maximum of function (1), which is V. Thus, regardless of whether the variance constraint is an equality constraint or an inequality constraint, the optimal solution will be the same in both of these cases.

Our approach to the optimum allocation Problem 1.1 will be twofold. First, in Section 3, we make use of the Karush–Kuhn–Tucker conditions (see Appendix B) to establish necessary and sufficient conditions, the so-called optimality conditions, for a solution to slightly reformulated optimization Problem 1.1, defined as a separate Problem 3.1. This task is one of the main objectives of this paper. Optimality conditions, which are often given as closed-form expressions, are fundamental to the analysis and development of effective algorithms for an optimization problem. Namely, algorithms recognize solutions by checking whether they satisfy various optimality conditions and terminate when such conditions hold. This elegant strategy has evident advantages over some alternative ad-hoc approaches, commonly used in survey sampling, which are usually tailored for a specific allocation algorithm being proposed. Next, in Section 4, we precisely define the *LRNA* algorithm which solves Problem 3.1 (and in consequence Problem 1.1) and based on the established optimality conditions we provide the formal proof of its optimality, which is the second main objective of this paper. To complement our work on this subject, we provide user-end function in R (see R Core Team, 2023) that implements the *LRNA*. This function is included in our package `stratallo` (Wójciak, 2023), which is published on the Comprehensive R Archive Network (CRAN) package repository.

2. Motivation

Optimum sample allocation Problem 1.1 is not only a theoretical problem, but it is also an issue of substantial practical importance. Usually, an increase in the number of samples entails greater costs of the data collection process. Thus, it is often demanded that total cost (2) be somehow minimized. On the other hand, the minimization of the cost should not cause significant reduction of the quality of the estimation, which can be measured by the variance (1). Hence, Problem 1.1 arises very naturally in the planning of sample surveys, when it is necessary to obtain an estimator $\hat{\theta}$ with some predetermined precision V that ensures the required level of estimation quality, while keeping the overall cost as small as possible. Problem 1.1 appears also in the context of optimum stratification and sample allocation between subpopulations in Skibicki and Wywiał (2002) or Lednicki and Wieczorkowski (2003). The authors of the latter paper incorporate variance equality constraint into the objective function and then use numerical algorithms (for minimization of a non-linear multivariate function) to find the minimum of the objective function. If the solution found violates any of the inequality constraints, then the objective function is properly adjusted and the algorithm is re-run again. See also a related paper by Wright, Noble and Bailer (2007), where the allocation under the constraint of the equal precision for estimation of the strata means was considered.

The problem of minimization of the total cost under constraint on stratified estimator's variance is well known in the domain literature. It was probably first formulated by Tore Dalenius in Dalenius (1949, 1953) and later in his Ph.D. thesis Dalenius (1957, Chapter 1.9, p. 19; Chapters 9.4 - 9.5, p. 199). Dalenius formed this allocation problem in the context of multicharacter (i.e. in the presence of several variables under study) stratified sampling (without replacement) and without taking into account upper-bounds constraints (5). He solved his problem with the use of simple geometric methods for the case of two strata and two estimated population means, indicating that the technique that was used is applicable also for the case with any number of strata and any number of variables. Among other resources that are worth mentioning are Yates (1960) and Chatterjee (1968).

Kokan and Khan (1967) considered a multicharacter generalization of Problem 1.1 and proposed a procedure that leads to the solution of this problem. The proof of the optimality of the obtained solution given by the authors is not strictly formal and, similarly to Dalenius' work, is based solely on geometrical methods. The *LRNA* algorithm presented in Section 4 below, can be viewed as a special case of that Kokan-Khan's procedure for a single study variable. It is this method that is generally accepted as the one that solves Problem 1.1 and is described in popular survey sampling textbooks such as, e.g. Särndal et al. (1992, Remark 12.7.1, p. 466) or Cochran (1977, Section 5.8, p. 104). For earlier references, see Hartley and Hocking (1963) or Kokan (1963), who discussed how to use non-linear programming technique to determine allocations in multicharacter generalization of Problem 1.1. More recent references can be made to Bethel (1989), who proposed a closed form expression (in terms of Lagrange multipliers) for a solution to relaxed Problem 1.1 without (5), as well as to Hughes and Rao (1979), who obtained the solution to Problem 1.1 by employing an extension of a result due to Thompson (1962). Eventually, for integer solution to Problem 1.1, we refer to Khan, Ahsan and Jahan (1997).

We would like to note that the form of the transformation (6) that we have chosen to convert Problem 1.1 into a convex optimization Problem 3.1 was not the only possible choice. An alternative transformation is for instance $z_h = \frac{1}{x_h}$, $h \in \mathcal{H}$, which was used by Kokan and Khan (1967) or Hughes and Rao (1979) in their approaches to (somewhat generalized) Problem 1.1. Nevertheless, as it turns out, transformation (6) causes that induced Problem 3.1 gains some interesting interpretation from the point of view of practical application. That is, if one treats z_h as stratum sample size x_h, $h \in \mathcal{H}$, then Problem 3.1 with $\widetilde{V} = n$ and $c_h = 1, h \in \mathcal{H}$, becomes a classical optimum allocation Problem A.1 with added one-sided lower-bounds constraints $z_h \geq m_h > 0$, $h \in \mathcal{H}$. Such allocation problem can be viewed as twinned to Problem A.2 and is itself interesting for practitioners. The lower bounds are necessary, e.g. for estimation of population strata variances S_h^2, $h \in \mathcal{H}$, which in practice are rarely known a priori. If they are to be estimated from the sample, it is required that at least $n_h \geq 2$, $h \in \mathcal{H}$. They also appear when one treats strata as domains and assigns upper bounds for variances of estimators of totals in domains. Such approach was considered, e.g. in Choudhry, Hidiroglou and Rao (2012), where the additional constraints $(\frac{1}{n_h} - \frac{1}{N_h})N_h^2 S_h^2 \leq R_h,\, h \in \mathcal{H}$, where $R_h,\, h \in \mathcal{H}$ are given constants, have been imposed. Obviously, this system of inequalities can be rewritten as lower-bounds constraints on n_h, i.e. $n_h \geq m_h = \frac{N_h^2 S_h^2}{R_h + N_h S_h^2},\, h \in \mathcal{H}$. The solution given in Choudhry et al. (2012) was obtained by the procedure based on the Newton-Raphson algorithm, a general-purpose root-finding numerical method. See also a related paper by Wright et al. (2007), where the allocation under the constraint of the equal precision for estimation of the strata means was considered. The affinity between allocation Problem 3.1 and Problem A.2 translates to significant similarities between the *LRNA* that solves Problem 3.1 and the popular recursive Neyman allocation procedure, *RNA*, that solves Problem A.2 (see Appendix A). To emphasize these similarities, the name *LRNA* was chosen for the former.

In summary, the *LRNA*, formulated as in this work, solves two different but related problems of optimum sample allocation that are of a significant practical importance, i.e. Problem 1.1 and Problem 3.1.

3. Optimality conditions

In this section, we establish a general form of the solution to (somewhat reformulated) Problem 1.1, the so-called optimality conditions. For this problem, the optimality conditions can be derived reliably from the Karush–Kuhn–Tucker (KKT) conditions, first derivative tests for a solution in nonlinear programming to be optimal (see Appendix B and the references given therein for more details). It is well known that for convex optimization problem (with some minor regularity conditions) the KKT conditions are not only necessary but also sufficient. Problem 1.1 is however not a convex optimization problem because the equality constraint function $\sum_{h\in\mathcal{H}} \frac{A_h^2}{x_h} - A_0 - V$ of $\mathbf{x} = (x_h,\, h \in \mathcal{H})$ is not affine and hence, the feasible set might not be convex. Nevertheless, it turns out that Problem 1.1 can be easily reformulated to a convex optimization Problem 3.1, by a simple change of its optimization variable from $\mathbf{x}$ to $\mathbf{z} = (z_h,\, \in \mathcal{H})$ with elements of the form:

$$z_h := \frac{A_h^2}{c_h x_h}, \qquad h \in \mathcal{H}. \tag{6}$$

Problem 3.1. Given a finite set $\mathcal{H} \neq \emptyset$ and numbers $A_h > 0$, $c_h > 0$, $m_h > 0$, $h \in \mathcal{H}$, $\widetilde{V} \geq \sum_{h\in\mathcal{H}} c_h m_h$,

$$\underset{\mathbf{z}=(z_h,\, h\in\mathcal{H})\in\mathbb{R}_+^{|\mathcal{H}|}}{\text{minimize}} \quad \sum_{h\in\mathcal{H}} \frac{A_h^2}{z_h} \tag{7}$$

$$\text{subject to} \quad \sum_{h\in\mathcal{H}} c_h z_h = \widetilde{V} \tag{8}$$

$$z_h \geq m_h, \qquad h \in \mathcal{H}. \tag{9}$$

For Problem 3.1 to be equivalent to Problem 1.1 under transformation (6), parameters

$m_h,\ h \in \mathcal{H}$, and $\widetilde{V}$ must be such that

$$\begin{aligned} m_h &:= \tfrac{A_h^2}{c_h M_h}, \qquad h \in \mathcal{H}, \\ \widetilde{V} &:= V + A_0 \geq A_0, \end{aligned} \tag{10}$$

where numbers $V,\ A_0,\ M_h,\ h \in \mathcal{H}$ are as in Problem 1.1. Nonetheless, as we explained at the end of Section 2 of this paper, Problem 3.1 can be considered as a separate allocation problem, unrelated to Problem 1.1; that is Problem A.1 with added one-sided lower-bounds constraints. For this reason, the only requirements imposed on these parameters are those given in the definition of Problem 3.1.

The auxiliary optimization Problem 3.1 is indeed a convex optimization problem as it is justified by Remark 3.1.

Remark 3.1. *Problem 3.1 is a convex optimization problem as its objective function* $f : \mathbb{R}_+^{|\mathcal{H}|} \to \mathbb{R}_+$,

$$f(\mathbf{z}) = \sum_{h \in \mathcal{H}} \tfrac{A_h^2}{z_h}, \tag{11}$$

and inequality constraint functions $g_h : \mathbb{R}_+^{|\mathcal{H}|} \to \mathbb{R}$,

$$g_h(\mathbf{z}) = m_h - z_h, \qquad h \in \mathcal{H}, \tag{12}$$

are convex functions, while the equality constraint function $w : \mathbb{R}_+^{|\mathcal{H}|} \to \mathbb{R}$,

$$w(\mathbf{z}) = \sum_{h \in \mathcal{H}} c_h z_h - \widetilde{V}$$

is affine. More specifically, Problem 3.1 is a convex optimization problem of a particular type in which inequality constraint functions (12) *are affine. See Appendix B for the definition of the convex optimization problem.*

As we shall see in Theorem 3.1, the optimization Problem 3.1 has a unique optimal solution. Consequently, due to transformation (6) and given (10), vector

$$\mathbf{x}^* = \left(\tfrac{A_h^2}{c_h z_h^*},\ h \in \mathcal{H} \right) \tag{13}$$

is a unique optimal solution of Problem 1.1, where $\mathbf{z}^* = (z_h^*,\ h \in \mathcal{H})$ is a solution to Problem 3.1. For this reason, for the remaining part of this work, our focus will be on a solution to Problem 3.1. We also note here that the solution to Problem 3.1 is trivial in the case of $\widetilde{V} = \sum_{h \in \mathcal{H}} c_h m_h$, i.e.: $\mathbf{z}^* = (m_h,\ h \in \mathcal{H})$.

Before we establish necessary and sufficient optimality conditions for a solution to convex optimization Problem 3.1, we first define a set function s, which considerably simplifies notation and many calculations that are carried out in this and subsequent section.

Definition 3.1. Let $\mathcal{H}$, A_h, c_h, m_h, $h \in \mathcal{H}$, and $\widetilde{V}$ be as in Problem 3.1. Set function s is defined as:

$$s(\mathcal{L}) = \frac{\widetilde{V} - \sum_{h\in\mathcal{L}} c_h m_h}{\sum_{h\in\mathcal{H}\setminus\mathcal{L}} A_h\sqrt{c_h}}, \qquad \mathcal{L} \subsetneq \mathcal{H}. \tag{14}$$

Below, we will introduce the notation of vector $\mathbf{z}^{\mathcal{L}} = (z_h^{\mathcal{L}},\, h \in \mathcal{H})$. It turns out that the solution to Problem 3.1 is necessarily of the form (15) with the set $\mathcal{L} \subseteq \mathcal{H}$ defined implicitly through the inequality of a certain form given in Theorem 3.1.

Definition 3.2. Let $\mathcal{H}$, A_h, c_h, m_h, $h \in \mathcal{H}$, $\widetilde{V}$ be as in Problem 3.1 and let $\mathcal{L} \subseteq \mathcal{H}$. Vector $\mathbf{z}^{\mathcal{L}} = (z_h^{\mathcal{L}},\, h \in \mathcal{H})$ is defined as follows

$$z_h^{\mathcal{L}} = \begin{cases} m_h, & h \in \mathcal{L} \\ \frac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}) & h \in \mathcal{H}\setminus\mathcal{L}. \end{cases} \tag{15}$$

The following Theorem 3.1 characterizes the form of the optimal solution to Problem 3.1 and therefore is the key theorem of this paper.

Theorem 3.1 (Optimality conditions)**.** *The optimization Problem 3.1 has a unique optimal solution. Point* $\mathbf{z}^* = (z_h^*,\, h \in \mathcal{H}) \in \mathbb{R}_+^{|\mathcal{H}|}$ *is a solution to optimization Problem 3.1 if and only if* $\mathbf{z}^* = \mathbf{z}^{\mathcal{L}^*}$ *with* $\mathcal{L}^* \subseteq \mathcal{H}$*, such that one of the following two cases holds:*

CASE I: $\mathcal{L}^* \subsetneq \mathcal{H}$ *and*

$$\mathcal{L}^* = \left\{ h \in \mathcal{H} :\, s(\mathcal{L}^*) \le \frac{\sqrt{c_h} m_h}{A_h} \right\}, \tag{16}$$

where set function s *is defined in* (14).

CASE II: $\mathcal{L}^* = \mathcal{H}$ *and*

$$\widetilde{V} = \sum_{h\in\mathcal{H}} c_h m_h. \tag{17}$$

Proof. We first prove that the solution to Problem 3.1 exists and it is unique. In the optimization Problem 3.1, a feasible set $F := \{\mathbf{z} \in \mathbb{R}_+^{|\mathcal{H}|} :\, \text{(8) and (9) are satisfied}\}$ is non-empty as guaranteed by the requirement $\widetilde{V} \ge \sum_{h\in\mathcal{H}} c_h m_h$. The objective function in (7) attains its minimum on F since it is a continuous function on F and F is closed and bounded. Finally, the uniqueness of the solution is due to strict convexity of the objective function on the set F.

As mentioned at the beginning of Section 3, the form of the solution to Problem 3.1, can be derived from the KKT conditions (see Appendix B). Following the notation of Remark 3.1, gradients of the objective function f and constraint functions w, g_h, $h \in \mathcal{H}$, are as follows:

$$\nabla f(\mathbf{z}) = \left(-\frac{A_h^2}{z_h^2},\, h \in \mathcal{H}\right), \quad \nabla w(\mathbf{z}) = (c_h,\, h \in \mathcal{H}), \quad \nabla g_h(\mathbf{z}) = -\underline{1}_h, \quad \mathbf{z} \in \mathbb{R}_+^{|\mathcal{H}|},$$

where $\underline{1}_h$ is a vector with all entries 0 except the entry at index h, which is 1. Consequently, the KKT conditions (33) assume the following form for the optimization Problem 3.1:

$$-\frac{A_h^2}{z_h^{*2}} + \lambda c_h - \mu_h = 0, \qquad h \in \mathcal{H}, \tag{18}$$

$$\sum_{h \in \mathcal{H}} c_h z_h^* - \widetilde{V} = 0, \tag{19}$$

$$m_h - z_h^* \leq 0, \qquad h \in \mathcal{H}, \tag{20}$$

$$\mu_h (m_h - z_h^*) = 0, \qquad h \in \mathcal{H}. \tag{21}$$

Following Theorem B.1 and Remark 3.1, in order to prove Theorem 3.1, it suffices to show that there exist $\lambda \in \mathbb{R}$ and $\mu_h \geq 0,\ h \in \mathcal{H}$, such that (18) - (21) are met for $\mathbf{z}^* = \mathbf{z}^{\mathcal{L}^*}$ with $\mathcal{L}^* \subseteq \mathcal{H}$ satisfying conditions of CASE I or CASE II.

CASE I: Following (15) and (14), we get

$$\sum_{h \in \mathcal{H}} c_h z_h^* = \sum_{h \in \mathcal{L}^*} c_h m_h + \sum_{h \in \mathcal{H} \setminus \mathcal{L}^*} c_h \frac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}^*) = \widetilde{V},$$

and hence, the condition (19) is always satisfied. Let $\lambda = \frac{1}{s^2(\mathcal{L}^*)}$, where $s(\mathcal{L}^*) > 0$ is defined in (14), and

$$\mu_h = \begin{cases} \lambda c_h - \frac{A_h^2}{m_h^2}, & h \in \mathcal{L}^* \\ 0, & h \in \mathcal{H} \setminus \mathcal{L}^*. \end{cases} \tag{22}$$

Note that $\mu_h \geq 0,\ h \in \mathcal{L}^*$, due to (16). Then, the condition (18) is clearly satisfied. Inequalities (20) and equalities (21) are trivial for $h \in \mathcal{L}^*$ since $z_h^* = m_h$. For $h \in \mathcal{H} \setminus \mathcal{L}^*$, inequalities (20) follow from (16), i.e. $\frac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}^*) > m_h$, whilst (21) hold true due to $\mu_h = 0$.

CASE II: Take arbitrary $\lambda \geq \max_{h \in \mathcal{H}} \frac{A_h^2}{m_h^2 c_h}$ and $\mu_h = \lambda c_h - \frac{A_h^2}{m_h^2},\ h \in \mathcal{H}$. Note that $\mu_h \geq 0,\ h \in \mathcal{H}$. Then, (18) - (21) are clearly satisfied for $(z_h^*,\ h \in \mathcal{H}) = (m_h,\ h \in \mathcal{H})$, whilst (19) follows after referring to (17).

□

Theorem 3.1 gives the general form of the optimum solution up to specification of the set $\mathcal{L}^* \subseteq \mathcal{H}$ that corresponds to the optimal solution $\mathbf{z}^* = \mathbf{z}^{\mathcal{L}^*}$. The issue of how to identify this set is the subject of the next section of this paper.

4. Recursive Neyman algorithm under lower-bounds constraints

In this section, we formalize the definition of the existing algorithm, termed here *LRNA*, solving Problem 3.1 and provide a formal proof of its optimality. The proof given is based on the optimality conditions formulated in Theorem 3.1.

Algorithm *LRNA*

Input: $\mathcal{H},\ (A_h)_{h\in\mathcal{H}},\ (c_h)_{h\in\mathcal{H}},\ (m_h)_{h\in\mathcal{H}},\ \widetilde{V}$.

Require: $A_h > 0,\ c_h > 0,\ m_h > 0,\ h \in \mathcal{H},\ \widetilde{V} \geq \sum_{h\in\mathcal{H}} c_h m_h$.

Step 1: Let $\mathcal{L} = \emptyset$.

Step 2: Determine $\widetilde{\mathcal{L}} = \left\{h \in \mathcal{H} \setminus \mathcal{L} :\ \frac{A_h}{\sqrt{c_h}} s(\mathcal{L}) \leq m_h\right\}$, where function s is defined in (14).

Step 3: If $\widetilde{\mathcal{L}} = \emptyset$, go to Step 4 . Otherwise, update $\mathcal{L} \leftarrow \mathcal{L} \cup \widetilde{\mathcal{L}}$ and go to Step 2 .

Step 4: Return $\mathbf{z}^* = (z_h^*,\ h \in \mathcal{H})$ with $z_h^* = \begin{cases} m_h, & h \in \mathcal{L} \\ \frac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}), & h \in \mathcal{H} \setminus \mathcal{L}. \end{cases}$

Theorem 4.1. *The* LRNA *provides an optimal solution to optimization Problem 3.1.*

Before we prove Theorem 4.1, we first reveal certain monotonicity property of set function s, defined in (14), that will be essential to the proof of this theorem.

Lemma 4.2. *Let* $\mathcal{A} \subseteq \mathcal{B} \subsetneq \mathcal{H}$. *Then*

$$s(\mathcal{A}) \geq s(\mathcal{B}) \quad \Leftrightarrow \quad s(\mathcal{A}) \sum_{h\in\mathcal{B}\setminus\mathcal{A}} A_h\sqrt{c_h} \leq \sum_{h\in\mathcal{B}\setminus\mathcal{A}} c_h m_h, \tag{23}$$

where set function s *is defined in* (14).

Proof. Clearly, for any $\alpha \in \mathbb{R},\ \beta \in \mathbb{R}, \delta \in \mathbb{R},\ \gamma \in \mathbb{R}_+$, such that $\gamma + \delta > 0$, we have

$$\tfrac{\alpha+\beta}{\gamma+\delta} \geq \tfrac{\alpha}{\gamma} \quad \Leftrightarrow \quad \tfrac{\alpha+\beta}{\gamma+\delta}\delta \leq \beta. \tag{24}$$

To prove (23), take

$$\alpha = \widetilde{V} - \sum_{h\in\mathcal{B}} c_h m_h \qquad \beta = \sum_{h\in\mathcal{B}\setminus\mathcal{A}} c_h m_h$$

$$\gamma = \sum_{h\in\mathcal{H}\setminus\mathcal{B}} A_h\sqrt{c_h} \qquad \delta = \sum_{h\in\mathcal{B}\setminus\mathcal{A}} A_h\sqrt{c_h}.$$

Then, $\frac{\alpha}{\gamma} = s(\mathcal{B}),\ \frac{\alpha+\beta}{\gamma+\delta} = s(\mathcal{A})$, and hence (23) holds as an immediate consequence of (24). ☐

We are now ready to give the proof of Theorem 4.1.

Proof of Theorem 4.1. Let $\mathcal{L}_r,\ \widetilde{\mathcal{L}}_r$ denote sets $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ respectively, as in the r-th iteration of the *LRNA* algorithm, at the moment after Step 2 and before Step 3. The iteration index r takes on values from set $\{1, \ldots, r^*\}$, where $r^* \geq 1$ indicates the final iteration of the algorithm. Under this notation, we have $\mathcal{L}_1 = \emptyset$ and in general for subsequent iterations, if any (i.e. if $r^* \geq 2$), we get

$$\mathcal{L}_r = \mathcal{L}_{r-1} \cup \widetilde{\mathcal{L}}_{r-1} = \bigcup_{i=1}^{r-1} \widetilde{\mathcal{L}}_i, \qquad r = 2, \ldots, r^*. \tag{25}$$

To prove Theorem 4.1, we have to show that:

(I) the algorithm terminates in a finite number of iterations, i.e. $r^* < \infty$,

(II) the solution computed at r^* is optimal.

The proof of (I) is relatively straightforward. In every iteration $r = 2, \dots, r^* \geq 2$, the domain of discourse for $\widetilde{\mathcal{L}}_r$ at Step 2 is $\mathcal{H} \setminus \mathcal{L}_r = \mathcal{H} \setminus \bigcup_{i=1}^{r-1} \widetilde{\mathcal{L}}_i$, where $\widetilde{\mathcal{L}}_i \neq \emptyset$, $i = 1, \dots, r-1$. Therefore, in view of Step 3, we have that $r^* \leq |\mathcal{H}| + 1 < \infty$, where $r^* = |\mathcal{H}| + 1$ if and only if $|\widetilde{\mathcal{L}}_r| = 1$ for each $r = 1, \dots, r^* - 1$. In words, the algorithm terminates in at most $|\mathcal{H}| + 1$ iterations.

In order to prove (II), following Theorem 3.1, it suffices to show that for $\mathcal{L}_{r^*} \subsetneq \mathcal{H}$ (CASE I), for all $h \in \mathcal{H}$,

$$h \in \mathcal{L}_{r^*} \quad \Leftrightarrow \quad \tfrac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}_{r^*}) \leq m_h, \tag{26}$$

and for $\mathcal{L}_{r^*} = \mathcal{H}$ (CASE II):

$$\widetilde{V} = \sum_{h \in \mathcal{H}} c_h m_h. \tag{27}$$

We first note that the construction of the algorithm ensures that $\mathcal{L}_r \subsetneq \mathcal{H}$ for $r = 1, \dots, r^* - 1$, $r^* \geq 2$, and therefore $s(\mathcal{L}_r)$ for such r is well-defined.

CASE I. The $s(\mathcal{L}_{r^*})$ is well-defined since in this case $\mathcal{L}_{r^*} \subsetneq \mathcal{H}$.

Necessity: For $r^* = 1$, we have $\mathcal{L}_{r^*} = \emptyset$ and hence, the right-hand side of equivalence (26) is trivially met. Let $r^* \geq 2$. By Step 2 of the *LRNA*, we have

$$\tfrac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}_r) \leq m_h, \qquad h \in \widetilde{\mathcal{L}}_r, \tag{28}$$

for every $r = 1, \dots, r^* - 1$. Multiplying inequalities (28) sidewise by c_h and summing over $h \in \widetilde{\mathcal{L}}_r$, we get the right-hand side of equivalence (23) with $\mathcal{A} = \mathcal{L}_r$ and $\mathcal{B} = \mathcal{L}_r \cup \widetilde{\mathcal{L}}_r = \mathcal{L}_{r+1} \subsetneq \mathcal{H}$. Then, by Lemma 4.2, the first inequality in (23) follows. Consequently,

$$s(\mathcal{L}_1) \geq \ldots \geq s(\mathcal{L}_{r^*}). \tag{29}$$

Now, assume that $h \in \mathcal{L}_{r^*} = \bigcup_{r=1}^{r^*-1} \widetilde{\mathcal{L}}_r$. Thus, $h \in \widetilde{\mathcal{L}}_r$ for some $r \in \{1, \dots, r^* - 1\}$, and again, using Step 2 of the *LRNA*, we get $\tfrac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}_r) \leq m_h$. Consequently, (29) yields $\tfrac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}_{r^*}) \leq m_h$.

Sufficiency: The proof is by establishing a contradiction. Assume that $\tfrac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}_{r^*}) \leq m_h$ and $h \notin \mathcal{L}_{r^*}$. On the other hand, Step 3 of the *LRNA* yields $\tfrac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}_{r^*}) > m_h$ for $h \in \mathcal{H} \setminus \mathcal{L}_{r^*}$ (i.e. $h \notin \mathcal{L}_{r^*}$), which contradicts the assumption.

CASE II. Note that in this case it must be that $r^* \geq 2$. Following Step 2 of the *LRNA*, the only possibility is that for all $h \in \mathcal{H} \setminus \mathcal{L}_{r^*-1}$,

$$\tfrac{A_h}{\sqrt{c_h}}\, s(\mathcal{L}_{r^*-1}) = \tfrac{A_h}{\sqrt{c_h}} \frac{\widetilde{V} - \sum_{i \in \mathcal{L}_{r^*-1}} c_i m_i}{\sum_{i \in \mathcal{H} \setminus \mathcal{L}_{r^*-1}} A_i \sqrt{c_i}} \leq m_h. \tag{30}$$

Multiplying both sides of inequality (30) by c_h, summing it sidewise over $h \in \mathcal{H} \setminus \mathcal{L}_{r^*-1}$, we get $\widetilde{V} \leq \sum_{i \in \mathcal{H}} c_i m_i$, which, when combined with the requirement $\widetilde{V} \geq \sum_{h \in \mathcal{H}} c_h m_h$, yields $\widetilde{V} = \sum_{h \in \mathcal{H}} c_h m_h$, i.e. (27).

□

Following Theorem 4.1 and given (13), the optimal solution to Problem 1.1 is vector $\mathbf{x}^* = (x_h^*,\, h \in \mathcal{H})$ with elements of the form:

$$x_h^* = \begin{cases} M_h, & h \in \mathcal{L}^* \\ \frac{A_h}{\sqrt{c_h}} \frac{\sum_{i \in \mathcal{H} \setminus \mathcal{L}^*} A_i \sqrt{c_i}}{V + A_0 - \sum_{i \in \mathcal{L}^*} \frac{A_i^2}{M_i}}, & h \in \mathcal{H} \setminus \mathcal{L}^*, \end{cases} \tag{31}$$

where $\mathcal{L}^* \subseteq \mathcal{H}$ is determined by the *LRNA*.

5. Final remarks and conclusions

Within this work we formulated the optimality conditions for an important problem of minimum cost allocation under constraints on stratified estimator's variance and maximum samples sizes in strata. This allocation problem was defined in this paper as Problem 1.1 and converted to Problem 3.1 through transformation (6) and under (10). Based on the established optimality conditions, we provided a formal and compact proof of the optimality of the *LRNA* algorithm that solves the allocation problem mentioned. As already outlined at the end of Section 2 of this paper, Problem 3.1 can be viewed in two ways, each of which is of a great practical importance. That is, apart from its primary interpretation as the problem of minimizing the total cost under given constraints, it can also be perceived as the problem of minimizing the stratified estimator's variance under constraint on total sample size (i.e. Problem A.1) and constraints imposed on minimum sample sizes in the strata. For this reason, all the results of this work established in relation to Problem 3.1 (i.e. optimality conditions and the *LRNA*) are of such a twofold nature.

For the reasons mentioned at the end of Section 2, the *LRNA* can be considered as a counterparty to the *RNA*. This resemblance is particularly desirable, given the popularity, simplicity as well as relatively high computational efficiency of the latter algorithm. Among the alternative approaches that could potentially be adapted to solve Problem 3.1 are the ideas that underlay the existing algorithms dedicated to Problem A.2, i.e.: *SGA* (Stenger and Gabler, 2005, Wesołowski, Wieczorkowski and Wójciak, 2021) and *COMA* (Wesołowski et al., 2021). For integer-valued algorithms dedicated to Problem 3.1 with added upper-bounds constraints, see Friedrich, Münnich, de Vries and Wagner (2015), Wright (2017, 2020). Nevertheless, it should be noted that integer-valued algorithms are typically relatively slow compared to not-necessarily integer-valued algorithms. As pointed out in Friedrich et al. (2015), computational efficiency of integer-valued allocation algorithms becomes an issue for cases with "many strata or when the optimal allocation has to be applied repeatedly, such as in iterative solutions of stratification problems".

Finally, we would like to emphasize that the optimality conditions established in Theorem 3.1 can be used as a baseline for the development of new algorithms that provide solution to the optimum allocation problem considered in this paper. For instance, such algorithms could be derived by exploiting the ideas embodied in *SGA* or *COMA*, dedicated to Problem A.2 as indicated above.

Theoretical results obtained in this paper are complemented by the R-implementation (R Core Team, 2023) of the *LRNA*, which we include in our publicly available package `stratallo` (Wójciak, 2023).

Acknowledgements

I am very grateful to Jacek Wesołowski from Warsaw University of Technology, my research supervisor, for his patient guidance on this research work. Many thanks to Robert Wieczorkowski from Statistics Poland for advice and explanations on the topic of optimum stratification. I would also like to thank Reviewers for taking the necessary time and effort to review the manuscript. In particular, I express my gratitude to the second of the Reviewers for his expertise, valuable suggestions and for pointing to the existing papers, particularly important from the point of view of the subject I am addressing in this work.

References

Bethel, J., (1989). Sample allocation in multivariate surveys. *Survey Methodology*, 15(1), pp. 47–57. https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X198900114578

Boyd, S., and Vandenberghe, L., (2004). *Convex Optimization*, Cambridge University Press, Cambridge.

Chatterjee, S., (1968). Multivariate Stratified Surveys. *Journal of the American Statistical Association*, 63(322), pp. 530–534. https://doi.org/10.2307/2284023

Choudhry, G. H., Hidiroglou, M. and Rao, J., (2012). On sample allocation for efficient domain estimation. *Survey Methodology*, 38(1), pp. 23–29. https://www150.statcan.gc.ca/n1/en/catalogue/12-001-X201200111682

Cochran, W. G., (1977). *Sampling Techniques*, 3rd edn, John Wiley & Sons, New York.

Dalenius, T., (1949). Den nyare utvecklingen inom teorin och metodiken för stickprovsundersökningar. *Förhandlingar vid Nordiska Statistikermötet i Helsingfors den 13 och 14 juni 1949*, Helsingfors, pp. 46–74.

Dalenius, T., (1953). The multivariate sampling problem. *Skandinavisk Aktuarietidskrift*, 36, pp. 92–102.

Dalenius, T., (1957). *Sampling in Sweden: Contributions to the Methods and Theories of Sample Survey Practice*, Almqvist & Wiksell, Stockholm.

Friedrich, U., Münnich, R., de Vries, S. and Wagner, M., (2015). Fast integer-valued algorithms for optimal allocations under constraints in stratified sampling. *Computational Statistics & Data Analysis*, 92, pp. 1–12. https://www.sciencedirect.com/science/article/pii/S0167947315001413

Hartley, H. O. and Hocking, R. R., (1963). Convex Programming by Tangential Approximation. *Management Science*, 9(4), pp. 600–612. https://doi.org/10.1287/mnsc.9.4.600

Hughes, E., and Rao, J. N. K., (1979). Some problems of optimal allocation in sample surveys involving inequality constraints. *Communications in Statistics-theory and Methods*, 8, pp. 1551–1574.

Khan, M. G. M., Ahsan, M. J. and Jahan, N., (1997). Compromise allocation in multivariate stratified sampling: An integer solution. *Naval Research Logistics (NRL)*, 44(1), pp. 69–79. https://doi.org/10.1002/(SICI)1520-6750(199702)44:1<69::AID-NAV4>3.0.CO;2-K

Kokan, A. R., (1963). Optimum Allocation in Multivariate Surveys. *Journal of the Royal Statistical Society. Series A (General)*, 126(4), pp. 557–565. https://doi.org/10.2307/2982579

Kokan, A. R. and Khan, S., (1967). Optimum Allocation in Multivariate Surveys: An Analytical Solution, *Journal of the Royal Statistical Society. Series B (Methodological)*, 29(1), pp. 115–125. https://doi.org/10.1111/j.2517-6161.1967.tb00679.x

Lednicki, B. and Wieczorkowski, R., (2003). Optimal Stratification and Sample Allocation between Subpopulations and Strata. *Statistics in Transition*, 6(2), pp. 287–305. https://stat.gov.pl/download/gfx/portalinformacyjny/en/defaultstronaopisowa/3432/1/1/sit_volume_4-7.zip

Neyman, J., (1934). On the Two Different Aspects of the Representative Method: the Method of Stratified Sampling and the Method of Purposive Selection. *Journal of the Royal Statistical Society*, 97(4), pp. 558–625.

R Core Team, (2023). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Skibicki, M. and Wywiał, J., (2002). On optimal sample allocation in strata, *in* J. Paradysz (ed.). *Statystyka regionalna w służbie samorządu lokalnego i biznesu*, Internetowa Oficyna Wydawnicza Centrum Statystyki Regionalnej, Poznań, pp. 29–37.

Stenger, H. and Gabler, S., (2005). Combining random sampling and census strategies - Justification of inclusion probabilities equal to 1. *Metrika*, 61(2), pp. 137–156. https://doi.org/10.1007/s001840400328

Särndal, C.-E., Swensson, B. and Wretman, J., (1992). *Model Assisted Survey Sampling*, Springer, New York.

Tschuprow, A. A., (1923a). On the mathematical expectation of the moments of frequency distributions in the case of correlated observation (Chapters 1-3), *Metron*, 2(3), pp. 461–493.

Tschuprow, A. A., (1923b). On the mathematical expectation of the moments of frequency distributions in the case of correlated observation (Chapters 4-6), *Metron*, 2(4), pp. 636–680.

Thompson, W. A., (1962). The problem of negative estimates of variance components. *Annals of Mathematical Statistics*, 33, pp. 273–289.

Wesołowski, J., Wieczorkowski, R. and Wójciak, W., (2021). Optimality of the Recursive Neyman Allocation. *Journal of Survey Statistics and Methodology*, 10(5), pp. 1263–1275. https://academic.oup.com/jssam/article-pdf/10/5/1263/46878255/smab018.pdf

Wright, S. E., Noble, R. and Bailer, A. J., (2007). Equal-precision allocations and other constraints in stratified random sampling. *Journal of Statistical Computation and Simulation*, 77(12), pp. 1081–1089. https://doi.org/10.1080/10629360600897191

Wright, T., (2017). Exact optimal sample allocation: More efficient than Neyman. *Statistics & Probability Letters*, 129, pp. 50–57. https://www.sciencedirect.com/science/article/pii/S0167715217301657

Wright, T., (2020). A general exact optimal sample allocation algorithm: With bounded cost and bounded sample sizes. *Statistics & Probability Letters*, 165, pp. 108829. https://www.sciencedirect.com/science/article/pii/S0167715220301322

Wójciak, W., (2023). stratallo*: Optimum Sample Allocation in Stratified Sampling*. R package version 2.2.0. https://CRAN.R-project.org/package=stratallo

Yates, F., (1960). *Sampling Methods for Census and Surveys*, Griffin and Company, Ltd., London.

APPENDICES

A. Recursive Neyman allocation

The classical problem of optimum sample allocation is described e.g. in Särndal et al. (1992, Section 3.7.3, p. 104). It can be formulated in the language of mathematical optimization as Problem A.1.

Problem A.1. Given a finite set $\mathcal{H} \neq \emptyset$ and numbers $A_h > 0,\ h \in \mathcal{H},\ 0 < n \leq N$,

$$\begin{aligned} \underset{\mathbf{x} = (x_h,\, h \in \mathcal{H}) \in \mathbb{R}_+^{|\mathcal{H}|}}{\text{minimize}} \quad & \sum_{h \in \mathcal{H}} \frac{A_h^2}{x_h} \\ \text{subject to} \quad & \sum_{h \in \mathcal{H}} x_h = n. \end{aligned}$$

The solution to Problem A.1 is $\mathbf{x}^* = (x_h^*,\ h \in \mathcal{H})$ with elements of the form:

$$x_h^* = A_h \frac{n}{\sum_{i \in \mathcal{H}} A_i}, \qquad h \in \mathcal{H}.$$

It was established by Tschuprow (1923a,b) and Neyman (1934) for *stratified π estimator* of the population total with *simple random sampling without replacement* design in strata, in the case of which $A_h = N_h S_h,\ h \in \mathcal{H}$, where S_h denotes stratum standard deviation of a given study variable. See, e.g. Särndal et al. (1992, Section 3.7.4.i., p. 106) for more details.

The recursive Neyman allocation algorithm, denoted here as *RNA*, is a well-established allocation procedure that finds a solution to the classical optimum sample allocation Problem A.1 with added one-sided upper-bounds constraints, defined here as Problem A.2.

Problem A.2. Given a finite set $\mathcal{H} \neq \emptyset$ and numbers $A_h > 0,\ M_h > 0$, such that $M_h \leq N_h,\ h \in \mathcal{H}$, and $0 < n \leq \sum_{h \in \mathcal{H}} M_h$,

$$\begin{aligned} \underset{\mathbf{x} = (x_h,\, h \in \mathcal{H}) \in \mathbb{R}_+^{|\mathcal{H}|}}{\text{minimize}} \quad & \sum_{h \in \mathcal{H}} \frac{A_h^2}{x_h} \\ \text{subject to} \quad & \sum_{h \in \mathcal{H}} x_h = n \\ & x_h \leq M_h, \qquad h \in \mathcal{H}. \end{aligned}$$

Algorithm *RNA*

Input: $\mathcal{H}$, $(A_h)_{h\in\mathcal{H}}$, $(M_h)_{h\in\mathcal{H}}$, n.

Require: $A_h > 0$, $M_h > 0$, $h \in \mathcal{H}$, $0 < n \leq \sum_{h\in\mathcal{H}} M_h$.

Step 1: Let $\mathcal{U} = \emptyset$.

Step 2: Determine $\widetilde{\mathcal{U}} = \left\{h \in \mathcal{H} \setminus \mathcal{U} :\ A_h \frac{n-\sum_{i\in\mathcal{U}} M_i}{\sum_{i\in\mathcal{H}\setminus\mathcal{U}} A_i} \geq M_h\right\}$.

Step 3: If $\widetilde{\mathcal{U}} = \emptyset$, go to Step 4 . Otherwise, update $\mathcal{U} \leftarrow \mathcal{U} \cup \widetilde{\mathcal{U}}$, and go to Step 2 .

Step 4: Return $\mathbf{x}^* = (x_h^*,\ h \in \mathcal{H})$ with $x_h^* = \begin{cases} M_h, & h \in \mathcal{U} \\ A_h \frac{n-\sum_{i\in\mathcal{U}} M_i}{\sum_{i\in\mathcal{H}\setminus\mathcal{U}} A_i}, & h \in \mathcal{H} \setminus \mathcal{U}. \end{cases}$

For more information on this recursive procedure see Särndal et al. (1992, Remark 12.7.1, p. 466) and Wesołowski et al. (2021) for the proof of its optimality.

B. Convex optimization scheme and the KKT conditions

A convex optimization problem is an optimization problem in which the objective function is a convex function and the feasible set is a convex set. In standard form it is written as

$$\begin{aligned} \underset{\mathbf{x}\in\mathcal{D}}{\text{minimize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & w_i(\mathbf{x}) = 0, \quad i = 1, \ldots, k \\ & g_j(\mathbf{x}) \leq 0, \quad j = 1, \ldots, \ell, \end{aligned} \tag{32}$$

where $\mathcal{D} \subseteq \mathbb{R}^p$, $p \in \mathbb{N}_+$, the objective function $f : \mathcal{D}_f \subseteq \mathbb{R}^p \to \mathbb{R}$ and inequality constraint functions $g_j : \mathcal{D}_{g_j} \subseteq \mathbb{R}^p \to \mathbb{R}$, $j = 1, \ldots, \ell$, are convex, whilst equality constraint functions $w_i : \mathcal{D}_{w_i} \subseteq \mathbb{R}^p \to \mathbb{R}$, $i = 1, \ldots, k$, are affine. Here, $\mathcal{D} = \mathcal{D}_f \cap \bigcap_{i=1}^{k} \mathcal{D}_{w_i} \cap \bigcap_{j=1}^{\ell} \mathcal{D}_{g_j}$ denotes a common domain of all the functions. Point $\mathbf{x} \in \mathcal{D}$ is called *feasible* if it satisfies all of the constraints, otherwise the point is called *infeasible*. An optimization problem is called *feasible* if there exists $\mathbf{x} \in \mathcal{D}$ that is *feasible*, otherwise the problem is called *infeasible*.

In the context of the optimum allocation Problem 3.1 discussed in this paper, we are interested in a particular type of the convex problem, i.e. (32) in which all inequality constraint functions g_j, $j = 1, \ldots, \ell$, are affine. It is well known, see, e.g. the monograph Boyd and Vandenberghe (2004), that the solution for such an optimization problem can be identified through the set of equations and inequalities known as the Karush-Kuhn-Tucker (KKT) conditions, which in this case are not only necessary but also sufficient.

Theorem B.1 (KKT conditions for convex optimization problem with affine inequality constraints)**.** *A point* $\mathbf{x}^* \in \mathcal{D} \subseteq \mathbb{R}^p$ *is a solution to the convex optimization problem* (32) *in which functions* g_j, $j = 1, \ldots, \ell$, *are affine if and only if there exist numbers* $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, k$, *and* $\mu_j \geq 0$, $j = 1, \ldots, \ell$, *called KKT multipliers, such that*

$$\begin{aligned} \nabla f(\mathbf{x}^*) + \sum_{i=1}^{k} \lambda_i \nabla w_i(\mathbf{x}^*) + \sum_{j=1}^{\ell} \mu_j \nabla g_j(\mathbf{x}^*) = \mathbf{0} & \\ w_i(\mathbf{x}^*) = 0, \qquad i = 1, \ldots, k & \\ g_j(\mathbf{x}^*) \leq 0, \qquad j = 1, \ldots, \ell & \\ \mu_j g_j(\mathbf{x}^*) = 0, \qquad j = 1, \ldots, \ell. & \end{aligned} \tag{33}$$