

POVZETEK

V prvem poglavju opišemo splošne lastnosti in matematično formulacijo linearnega programiranja. V uvodnih primerih prikažemo, kako prepoznamo nalogo in določimo problem za nadaljnje reševanje. Omejimo se na simpleksno metodo, ki poišče rešitev naloge linearnega programiranja. Delovanje te metode in posebnosti, na katere naletimo pri reševanju teh nalog, prikažemo na primerih. Poglavje zaključimo s teorijo o dualnosti, ki pogosto olajša reševanje originalne naloge.

V drugem poglavju se osredotočimo na uporabo linearnega programiranja v praksi, t.j. pri statistični zaščiti podatkov. Najprej opišemo kje v praksi naletimo na problem varovanja podatkov in kakšen je njegov pomen. Omejimo se na varovanje podatkov objavljenih v tabelah. Opišemo osnovne tipe spremenljivk, ki pogosto nastopajo pri oblikovanju tabel in igrajo zelo pomembno vlogo pri zaščiti. V nadaljevanju je opis različnih meril občutljivosti, ki določajo občutljivost celic. Za zaščito tabel uporabljamo različne metode. Natančneje opišemo metodi preoblikovanje tabel in zakrivanje. Po zaščiti preverimo ustreznost intervalov zakritih celic. Te intervale določimo z uporabo linearnega programiranja. Pogledamo, kako določimo tveganje razkritja objavljene tabele. Tveganje razkritja se nanaša na vrednostne in frekvenčne tabele. Pri objavi zaščiteneh podatkov je zelo pomembno, da objavljena tabela še vedno posreduje dovolj informacij uporabnikom. Opišemo nekaj meril za določitev izgube informacij. Pri zakrivanju podatkov so nam v pomoč različni programski paketi. Opišemo lastnosti programskega paketa τ -ARGUS in prikažemo njegovo delovanje na primeru, ki je vzet iz popisa prebivalstva.

Ključne besede:

linearno programiranje, simpleks algoritem, statistična zaščita podatkov, merilo občutljivosti, tveganje razkritja, izguba informacij.

ABSTRACT

Linear programming arises naturally in many optimization problems in economics and operations research. We present two examples that lead to a linear programming problems. The simplex algorithm which is the most widely used algorithm for solving such problems is presented with a few examples. In the end we briefly summarize the theory of duality which is important from both the theoretical and practical point of view.

The second chapter deals with statistical disclosure control. We first describe the legal framework for data protection and show how the requirements lead to mathematical problems. The discussion is limited to security of tabular data. We overview the types of variables that constitute the tables to be protected. Various sensitivity measures for the degree of protection are presented which enables us to identify cells in the tables whose publication would pose a disclosure risk. Among the techniques used to reduce the risk of unwanted disclosure the most common ones are table redesign and suppression. The first technique reduces the risk by lumping rows and/or columns in the table. The second simply suppresses the cell total. On one hand the published data should still contain enough information for researchers. On the other hand, however, it should not be possible to identify the values for individual units that contribute to cell totals. At this stage linear programming is used to compute feasibility intervals, i. e. the intervals where the values for individual units must lie given the information in the table. If the table is adequately protected such intervals must be broad enough but the protection has to lead to as small a loss of information as possible. There are various methods to quantify the loss of

Literatura

- [1] **V. Chvátal**: *Linear programming*: W. H. Freeman and Company, New York, 1983.
- [2] **L. Willenborg, T. de Waal**: *Statistical disclosure control in practice*: Springer-Verlag, New York, 1996.
- [3] **L. Willenborg, T. de Waal**: *Elements of statistical disclosure control*: Springer-Verlag, New York, 2001.
- [4] **A. Hundepool, A. van de Wetering**: τ -ARGUS, *User's Manual*,
[URL: <http://neon.vb.cbs.nl/casc/>], Netherlands, 2002.
- [5] **I. N. Bronstein, K. A. Semendjajew, G. Musiol, H. Mühlig**: *Primeri in grafične rešitve*, MATEMATIČNI PRIROČNIK: Tehniška založba Slovenije, 1997, str. 680, 681.
- [6] **K. Luhn**: *The determination of intervals of suppressed cells in an n - dimensional table*,
[URL: <http://neon.vb.cbs.nl/casc/>], Luxemburg, 2003.
- [7] **P. P. de Wolf**: *A heuristic approach to cell suppression in hierarchical tables*,
[URL: <http://neon.vb.cbs.nl/casc/>], Netherlands, 1999.
- [8] **J. A. Loeve**: *Notes on sensitivity measures and protection levels*,
[URL: <http://neon.vb.cbs.nl/casc/>], Netherlands, 2001.

Viri

1. Statistični urad Republike Slovenije, Popis prebivalstva, gospodinjstev in stanovanj 2002.