Nonlinear weighted least squares fitting Bass diffusion model

Dragan Jukić
Department of Mathematics, University J.J. Strossmayer of Osijek
Trg Ljudevita Gaja 6, HR-31000 Osijek, Croatia

Key words: *Bass model, least squares, total least squares, existence problem*

Abstract. The Bass model is one of the most well-known and widely used first-purchase diffusion models in marketing research. The main reason for this is that it finds its origin in a formal theory of product diffusion, and that the model parameters have an easy interpretation in terms of innovation and imitation. Namely, Bass classified adopters (first-time buyers) into two groups: innovators and imitators. Imitators, unlike innovators, are those buyers who are influenced in their adoption by the number of previous buyers.

Mathematically, the Bass diffusion model is described by the following differential equation:

$$\frac{dN(t)}{dt} = p[m - N(t)] + \frac{q}{m} N(t)[m - N(t)], \qquad N(0) = 0, \quad t \geq 0, \tag{1}$$

where $N(t)$ is the cumulative number of adopters of a new product at time t, parameter $m > 0$ is the total market potential for the new product, and the parameters $p > 0$ and $q \geq 0$ are the coefficients of innovation and imitation, respectively. The adoption rate, $\frac{dN(t)}{dt}$, is determined by two additive terms. The first term, $p[m - N(t)]$, represents adoptions due to innovators. The second term, $\frac{q}{m} N(t)[m - N(t)]$, represents adoptions due to imitators. The solution of (1) is given by

$$N(t; m, p, q) = m \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} e^{-(p+q)t}}, \qquad t \geq 0. \tag{2}$$

A number of procedures for estimating the parameters p, q and m of the Bass model have been suggested. Based on the analysis of some typical examples, the advantages and disadvantages of these approaches have been discussed by many authors.

We will present some of our results pertaining to the nonlinear weighted least squares (NWLS) and total least squares (TLS) fitting of the Bass curve (2) to the given data (w_i, t_i, N_i), $i = 1, \ldots, n$. Part $w_i > 0$ of the data stands for data weights. A necessary and sufficient condition which guarantees the existence of the NWLS and TLS estimate will be given. This condition is theoretical in nature and it is used for proving the remaining results. In the case when parameter m is bounded above by a known constant M, we will show that both NWLS and TLS estimate exist. Some generalization to the l_p norm $(1 \leq p < \infty)$ will be given.