

Multiplicative Marketing Mix Modeling simplified

Authors:

Deepen Garg | Sagar Shah |

Vaswati Ghosh | Saurabh Bagchi

"Statistical modeling beautifully represents reality in an objective quantitative way. Traditionally, Multiplicative MMM has the necessary framework but falls short of a business interpretation, which we have aimed to simplify"



Table of Contents

- 1. Introduction ----- 03
- 2. Comparing multiplicative model with linear regression model ----- 03
- 3. Deriving results in multiplicative log-linear model ----- 04
 - 3.1. Calculation of contribution to sales ----- 04
 - 3.2. Calculation of elasticity ----- 05
 - 3.3. Calculation of due-to ----- 05
 - 3.4. Missing value imputation for competitor data ----- 06
- 4. Conclusion ----- 06
- 5. References ----- 06
- 6. About Fractal Analytics ----- 07

1 Introduction

The optimal allocation of marketing funds has become an increasingly difficult problem across industries. Marketing Mix Models (MMM) quantify the contribution of marketing activities to sales with a view of calculating ROI, effectiveness and efficiency. We decompose the sales into base and incremental components. Base indicating consumer preferences (long-term) while incremental indicate short-term variations due to promotions, temporary price changes, and above the line media activity. Since models focus more on incrementality due to marketing activities, it falls short in calculating long-term impact of brand-building successful media campaigns and brand-eroding frequent promotions. However, even with these short-comings, Marketing Mix Modeling has proved to be an effective technique to allocate funds more analytically. (Cain, 2008, Tellis, 2006). Two commonly used methods include linear and multiplicative relationships between sales and marketing activities.

The multiplicative approach is more accurate but cumbersome. Some challenges arise in deriving:

- 1. Marketing effectiveness and ROI from log-linear marketing mix models.
- 2. Impact of competitor price in case of many missing data points is a challenge.

The solution we propose addresses both these issues.

2 Comparing multiplicative model with linear regression model

Multiplicative model’s benefits include synergistic impact of marketing activities, independence of shapes of transformations independent variables can take due to coefficient values and coefficients directly provide simple elasticity indications. Major limitations are inability to capture the S-shaped advertising curve appropriately, constant elasticity values. (Tellis, 2006).

| Variable Types | Product Type | Example | Linear Regression Transform | Multiplicative Regression Transform |
|----------------|--------------|--------------------------|-----------------------------|-------------------------------------|
| Base Variables | Own | Base Price/ Distribution | No transform | Log-transform |
| | Competitor | Base Price/ Distribution | No transform | Log-transform |
| Incremental | Own | Media Spends | Adstock | Adstock |
| | Competitor | Media Spends | Adstock | Adstock |

Log-linear regression is ‘theoretically’ better than linear regression because of its better physical explanation on the extreme values or boundary conditions of the regressors. In linear regression, price and distribution are treated the same as other incremental regressors, and even if we put price or distribution equal to zero; we still get a finite volume. This is contrary to common perception which dictates that volume should be zero or *infinite* if distribution or price is zero respectively. Log-linear regression takes care of this fact and with any ‘base-like’ variable equal to zero; it gives us a zero or *infinite* volume (depending upon the sign of associated coefficient). So, if we put x_k equal to zero, we get volume equal to either 0 or 0^{-1} that is *infinity*. Whereas, if we put any incremental regressor, y_k equal to zero, log-linear regression still gives us a finite volume, which is in line with our business perception which says that volume should not reduce to zero if no marketing or incremental activity is done.

Hence, log-linear regression provides a more robust theoretical explanation of the dynamics of the market while satisfying the boundary conditions.

3 Deriving results in multiplicative log-linear model

Multiplicative model presents challenges in deriving business results for contribution, elasticity and due to which can be solved with algebraic manipulations. The equations for the two types of regression models are as below:

- 1) Multivariate *Linear* regression :

$$\text{Volume} = a_1x_1 + a_2x_2 + b_1y_1 \dots + c$$
- 2) Multivariate Multiplicative (*log-linear*) regression:

$$\text{Volume} = K.x_1^{a_1}.x_2^{a_2}.x_3^{a_3} \dots e^{(b_1y_1 + b_2y_2 \dots)}$$

x_k : 'base-like' factors like price, distribution, competitors' price

y_k : incremental factors like own media, promotions, etc.,

a_k, b_k are variable coefficients, 'e' is Euler's constant

3.1 Calculation of contribution to sales

Challenge: Contribution is defined as the extra sales incurred due to the presence of that regressor. In a linear model, contribution is simply calculated by $a_k x_k$ for regressor x_k . In multiplicative log-linear modeling, we have a synergistic effect of the individual contributions. Since the volume is now construed as the result of the multiplication of various factors, it is not possible to distinguish the resulting sales into the separate non-overlapping effect of the different factors.

Solution: From first principles, contribution can be defined as the difference in total volume and the volume upon removal of the effect of that regressor. Let us take an example to understand this:

Let us assume the total volume is 12 and is due to 3 factors base, a and b . Base = 1; $a = 3$; $b = 4$

Now the contribution of a and b can be defined as the volume we would lose if we were to remove the effect of a or b . So if we remove the effect of a , we would get the volume of $12/3 = 4$. Hence the contribution of a is $12 - 4 = 8$.

Similarly upon removal of the effect of b , we are left with $12/4 = 3$ units which makes the contribution of b to be equal to 9 units.

Here, we would also like to define "Raw volume with respect to a regressor". In the preceding example, raw volume with respect to a is the volume we are left with upon removal of the effect of a which is 4. Raw volume for each regressor is defined as the hypothetical volume sold if the effect of that particular regressor is neutralized i.e. the beta of that regressor is assumed to be zero. It is the effective volume upon which the regressor acts and produces the actual sales as a result. This is not the final solution. We are still left with the problem of over-estimating the contributions from an additive standpoint. The contributions of base, a and b (1, 8 and 9 respectively) do not add up to 12. This is because we are trying to calculate contribution, which is a property of linear models, from multiplicative models. This problem can be solved through following approximation:

Suppose, the actual volume in a week is V_a and the total regressor contribution is V_p . So the difference ($V_a - V_p$) needs to be distributed among the various factors contributing to the volume. For simplicity, we assume we have three contributors (x_1, x_2 and x_3) with individual contribution as ' $-C_1$ ', ' C_2 ' and ' C_3 ' respectively. Here we are considering all the C_i 's to be positive numbers (>0) so the negative sign for $i=1$ emphasizes that it is a negative contribution.

Therefore scaled individual contribution will be as follows:

- $x_1: -C_1 + (C_1 * (V_a - V_p)) / (C_1 + C_2 + C_3)$
- $x_2: C_2 + (C_2 * (V_a - V_p)) / (C_1 + C_2 + C_3)$
- $x_3: C_3 + (C_3 * (V_a - V_p)) / (C_1 + C_2 + C_3)$

Hence,

Contribution (x_1) + Contribution (x_2) + Contribution (x_3) =

$$= -C_1 + C_2 + C_3 + V_a - V_p$$

$$= V_p + V_a - V_p \text{ (as by definition } V_p = -C_1 + C_2 + C_3)$$

$$= V_a$$

We have specifically taken up the example of a negative contributing factor to emphasize the fact that scaling should be done with respect to the absolute values of the contributors. This gives us an idea about the degree of influence each regressor has on the actual sales by time period

3.2 Calculation of Elasticity

Challenge: Elasticity can be defined as the percent change in total volumes with a unit percent change in the regressor. It can be thought of as $(dV/V)/(dx/x) = (dV/dx)*(x/V)$. For 'base-like' factors, the regressor elasticity is simply a_k , the coefficient. For 'incremental factors', it is equal to $\exp(b_k x_k) - 1$. For one period (week or month), it is fairly easy to define but for more than one observation, the value of x_k needs to be decided.

Solution: The best way to calculate elasticity would be to simulate using different values of regressors for each variable's elasticity. By increasing x_k by a tiny percent in all the periods and then calculate total % change in volumes. This reduces to using volume weighted average of all x_k for elasticity calculation in the formula $b_k x_k$.

3.3 Calculation of due-to

Challenge: The due-to of a regressor is the contribution of that regressor in total growth or decline in the volumes in time period t_2 compared to t_1 ($t_2 > t_1$). For linear modeling, it is the percentage change in total volumes exclusively due-to that regressor. So, it is calculated as difference in contribution of the regressor in time t_2 and time t_1 , divided by total volume sales of time t_1 . In log-linear modeling, we face the problem of decoupling the effect of other factors from the effect of the regressor in question. We cannot use the same formula of calculating due-to from contributions because of the effect of the factors on each other's contribution. It can be illustrated with following example:

Let us assume two time periods t_1 and t_2 with following values:

- $a=3, b=2$, So total volume = $1*a*b=6$ (at t_1) and contribution of a is approx. 3 and contribution of b is approx.. 2 (after scaling)
- $a=3, b=4$, So total volume = $1*a*b=12$ (at t_2) and contribution of a is approx. 5 and contribution of b is approx.. 6 (after scaling).

In this example, contribution of 'a' has increased due the effect of b but a is not the factor of growth here.

Solution: This solution defines "factors" for each regressor and measures the change in volumes in terms of change in factors. For example, in the previous example:

- % Change in volume: = $(12 - 6)/6 = 6/6 = 100\%$
growth = Factor of 2
- Change in factor of $a = 3/3 = 1$ which translates to 0% growth due to a
- Change in factor of $b = 4/2 = 2$ which translates to 100% growth due to b

Total change in factors of a and b = $1*2 = 2$
Which is why total volume at t_1 gets multiplied by 2 and we get the volume at t_2

This is how we decompose the total growth among individual regressors. To compare any two time periods, we need representative factors of each regressor. This can create a problem if we want to compare two years and we have data at weekly level. The calculation of a representative factor corresponding to x_k for each time period is then approximated by the sum of total volume, divided by the sum of raw volumes with respect to x_k in the same period. This will give us a representative factor for the period. Similarly, we can find a similar factor for another period and then find % change in the factor. After this exercise, we need to scale the due-to factors for all the concerned regressors to let them add up to the total growth observed. This scaling is again needed because we are approximating results from multiplicative model to depict the entities of additive model. The scaling method used is similar to the one we used while scaling contributions.

3.4 Missing value imputation for competitor data

Challenge: There are some time periods in the data when competitor volumes are zero due to no interceptable sales or because the competitor is just not present for that time period (maybe launched in the middle of our modeling period). We need to impute missing values here. This is aggravated especially if there are many such data points. If we put any finite value for the competitor price then we are implicitly assuming some impact of the competitor price on our volume even when competitor products are not selling, which can potentially destabilize the reliability of the learning on the competitor price.

Solution: Intuitively, if we use *infinite price* (or a very large value) to mark such occasions, we will be able to incorporate the correct impact in a log-linear model. Currently, we are using competitors' price as it is among 'base-like' factors in our log-linear model. This makes volumes go to *infinity* for infinite competitors' price. In order to rectify this problem, a transformation is needed which will satisfy the following essential constraints:

- Modeled volumes are finite at infinite competitor's price
- Modeled volumes are zero for zero competitor price
- The intercept normalizes input as well as output of the model

The proposed transformation is $(x/x+x')$ where x is the competitor's price, and x' is the mean of x across all real (excluding the very large value induced artificially) observations. So, the transformed equation will be: $\text{Volume} = K \cdot \text{Price}^{a1} \cdot \text{Dist}^{a2} \dots (x/x+x')^b$. This transformation takes care of all the constraints and presents a better solution than the usual method. It gives similar results to the usual method when used over continuous data. However, in the event of a lot of missing values, it gives a better approximation.

4 Conclusions

In this paper, we have proposed alternatives to calculate the results for Multiplicative MMM which encourage marketing levers having synergistic impacts on sales. Further, we have suggested an approach to enable a transformation of competitor price to correctly gauge its impact.

5 References

- Cain, P. M. (2008) 'Limitations of conventional marketing mix modeling' published in Admap
- Gerald J. Tellis, Chapter 24, Modeling Marketing Mix, University of Southern California

About Fractal Analytics

Fortune 500 companies recognize analytics is a competitive advantage to understand customers and make better decisions. Fractal Analytics delivers insight, innovation and impact to them through predictive analytics and visual story-telling.

Fractal Analytics' flagship Customer Genomics® solution helps marketers learn complex customer behaviour at an individual level. Its proprietary pattern recognition and machine-learning algorithms learn from every transaction and customer interaction, including social media, to help marketers build a complete view of individual customers across attitudinal and behavioural dimensions.

In June of 2013, global private equity firm TA Associates acquired a minority stake in the company for an investment of \$25 million, and in May of 2013, information technology and research advisor Gartner named Fractal as one of the a top five "Cool Vendors in Analytics."

For more information, contact us at:

+1 650 378 1284

info@fractalanalytics.com