# MMM Documentations – Methodologies

### Full Model Implementation

* Full implementation is available in VP [GitLab](https://gitlab.com/vistaprint-org/dna/customer-and-business-performance/measurement/marketingmixmodel)
* Please reach out CBPM if access needed

### VAR Model

#### Introduction

Vector autoregression (VAR) is a statistical model used to capture the relationship between multiple quantities as they change over time. VAR is a type of stochastic process model. VAR models generalize the single-variable (univariate) autoregressive model by allowing for multivariate time series. VAR models are often used in economics and the natural sciences.

The vector autoregression (VAR) is commonly used for forecasting systems of interrelated time series and for analyzing the dynamic impact of random disturbances on the system of variables. The reduced form VAR approach sidesteps the need for structural modeling by treating every endogenous variable in the system as a function of p-lagged values of all of the endogenous variables in the system.

#### Definition

We may write the stationary VAR(p) process as

Where

* is a vector endogenous variables,
* is a vector exogenous variables,
* are matrices of lag coefficients to be estimated,
* is a matrix of exogenous variable coefficients to be estimated,
* is a white noise innovation process, with

#### Estimation

To represent the model in a more compact way, let us assume

Thus,

Where

The Estimation are derived here –

### Impulse Response Function

#### Introduction

An impulse response function (IRF) of a time series model (or dynamic response of the system) measures the changes in the future responses of all variables in the system when a variable is shocked by an impulse.

#### Definition

In other words, the IRF at time t is the derivative of the responses at time t with respect to an innovation at time t0 (the time that innovation was shocked), t ≥ t0.

#### Theory

Denoting the known history of the economy up to time by the non-decreasing

information set , the generalized impulse response function of at horizon, advanced

in Koop et al. (1996), is defined by

#### Estimation

#### Here are basically two ways to estimate it. A traditional way using Cholesky Decomposition and a rather novel way called Generalized Impulse Response function.

**Method I - Cholesky**

As we know, VAR(p) can be rewritten as the infinite moving average representation if is covariance stationary. Thus,

And the coefficient matrices can be obtained using the following recursive relations:

Where

and for  and

Since an impulse response function measures the time profile of the effect of shocks at a given

point in time on the (expected) future values of variables in a dynamical system.

By combining equation with MA representation, we can have

This basically converting the problem the IRF estimation problem to a problem calculating both the and

The traditional approach, suggested by Sims (1980), is to resolve the problem surrounding the choice of by using the Cholesky decomposition of ,

where is an ***lower*** ***triangular*** matrix

Thus,

Denote that

, such that are orthogonalized. Hence, the vector of

the orthogonalized impulse response function of a unit shock to the equation on is

given by

Where is an selection vector with unity as its element and zeros elsewhere.

**Method II – Generalized Impulse Response**

An alternative approach would be to use equation directly, but instead of shocking all the elements of ; we could choose to shock only one element, say its element, and integrate out the effects of other shocks using an assumed or the historically observed distribution of the errors.

Assuming that has a multivariate normal distribution

Hence the (unscaled) generalized impulse response of the effect of a shock in the equation at time on is given by

### Forecast Error Variance Decomposition

#### Introduction

The variance decomposition indicates the amount of information each variable contributes to the other variables in the autoregression. It determines how much of the forecast error variance of each of the variables can be explained by exogenous shocks to the other variables. It is an extension of the Impulse Response results.

#### Definition

As mentioned, the above impulses can also be used in the derivation of the forecast error variance decompositions, defined as the proportion of the n-step ahead forecast error variance of variable which is accounted for by the innovations in variable in the VAR.

#### Estimation

According to the definition above, below are the estimation for FEVD for both IRF estimation

**Method I - Cholesky**

**Method II – Generalized Impulse Response**

### Confidence Interval Estimation

### ***Bootstrap methods***

Bootstrapping is any test or metric that uses random sampling with replacement and falls under the broader class of resampling methods. Bootstrapping assigns measures of accuracy (bias, variance, confidence intervals, prediction error, etc.) to sample estimates. This technique allows estimation of the sampling distribution of almost any statistic using random sampling methods.

Algorithm:

1. Obtain and =
2. Obtain via bootstrap and construct
3. Estimate using data constructed in 2).
4. Compute Upper bound and Lower bound of the IRF percentiles.

### ***Variation: bootstrap-after-the-bootstrap***

Algorithm:

1. Given obtain and construct
2. Estimate for each . If the bias is approximately constant in a neighbor of , .
3. Calculate the largest root of the system. If it is greater or equal than one, set - here the bias is irrelevant since estimates are superconsistent. Otherwise set .
4. Repeat 1) – 3) of Bootstrap algorithm, .

### Relative Elasticity and Optimal Budge Allocation

Given the estimated Elasticity, we can optimize the best budget allocation. Here we document the math behind it. Let us start with the following assumptions –

* Channel Spend
* Channel Elasticity
* Revenue, which is the target function

To maximize our target function, which is equivalent to maximize given , meaning we have a fixed budget A to spend.

To solve this optimization problem, basically we take derivative of both (1) and (2), thus arrive

,

This indicates that

Since , solving thus