

ANÁLISE FATORIAL: APLICAÇÃO EM DADOS DA INDÚSTRIA AUTOMOBILÍSTICA

José Milton Sanches*
Haroldo Santos Sanches**

RESUMO

O artigo tem por objetivo apresentar e comentar os dados de uma pesquisa econométrica visando a identificar informações relevantes sobre variáveis específicas da indústria automobilística brasileira no período 1990/2001. A partir da análise será possível identificar grupos de variáveis que se inter-relacionam em consonância com determinados fatores. Por outro lado, será possível identificar também que determinadas variáveis encontram-se isoladas em relação às outras. Para tanto, serão utilizadas ferramentas estatísticas disponibilizadas no SPSS (Statistical Package for the Social Sciences).

PALAVRAS-CHAVE: indústria automobilística, análise fatorial, desenvolvimento tecnológico.

ABSTRACT

This article has a purpose to accomplish an econometric research, to address the identification of relevant information about specific variables, related to the Brazilian car industry, during the period of 1990/2001. From this analysis it will be possible to identify groups of variables that interrelate in consonance with certain factors. On the other hand, it will also be possible to identify that certain variables are isolated in relation to others. For these, we shall use statistical tools made available in SPSS (Statistical Package for the Social Sciences).

KEY WORDS: car industry, factorial analysis, technological development.

Introdução

Até o início da última década do século passado, o Brasil apresentava um índice muito baixo de desenvolvimento tecnológico na indústria automobilística, com o mercado de importação de veículos automotores totalmente restrito a uma minoria de consumidores de alto poder aquisitivo. A produção nacional era dominada por um pequeno número de montadoras que atendiam à necessidade interna do país e exportavam para países menos desenvolvidos. Dessa forma, as montadoras de veículos existentes não se preocupavam com o desenvolvimento tecnológico de seus modelos.

Em 1990, o então recém-eleito Presidente da República, Fernando Collor de Melo, declarava que “os carros brasileiros eram verdadeiras carroças”. A frase de

* Mestre em economia pela PUC-SP e professor das Faculdades Padre Anchieta de Jundiá e da PUC-Campinas.
** Administrador formado pela PUC-Campinas.

forte impacto traduzia a necessidade de uma mudança radical neste mercado, que culminou com a liberação desenfreada das importações de veículos, num primeiro instante, e posteriormente, a instalação de novas montadoras no Brasil.

Desde então, as mudanças tecnológicas que foram implementadas contribuíram significativamente para o desenvolvimento do setor automobilístico, que passou a conquistar novos mercados, inclusive do primeiro mundo, dada a qualidade dos veículos aqui produzidos.

1. Base de dados

Os dados a serem utilizados na análise fatorial foram obtidos junto à ANFAVEA (Associação Nacional dos Fabricantes de Veículos Automotores). Para efeito de padronização, os dados originais das nove variáveis foram transformados em séries de números índices com base em 1990.

As variáveis a serem utilizadas na análise econométrica são as seguintes:

- Exportação total (em unidades);
- Exportação total (em valores);
- Vendas internas (em unidades);
- Produção total (em unidades);
- Número de empregados;
- Investimento total;
- Faturamento líquido total;
- Faturamento líquido do setor de Autopeças;
- Participação do setor automobilístico no PIB industrial.

~~Para efeito de identificação~~ Para efeito de identificação, os SPSS estão se adotando, respectivamente, os seguintes rótulos:

- Expunid;
- Expval;
- Viunid;
- Produnid;
- Empr;
- Invest;
- Fatliq;
- Fatap;
- Partpib.

As séries de números-índices (1990=100) para as nove variáveis estão apresentadas na Tabela I.

Tabela 1: Indicadores variados sobre o setor automobilístico



Fonte: ANFAVEA

Observe que no geral os indicadores apresentaram crescimento – ainda que com algumas oscilações – no período apresentado. A exceção fica por conta da variável “número de empregados” que visivelmente apresentou variação negativa no período.

2. Cálculos econométricos

Considerando que o objetivo do trabalho é estudar o inter-relacionamento entre um conjunto de variáveis observadas, utilizaremos a análise fatorial que é uma técnica estatística multivariada. Neste trabalho, o interesse é identificar a estrutura básica que está por trás de um conjunto de variáveis. Por exemplo, é possível reduzir um conjunto de 20 diferentes atributos – ou variáveis – medidos de um produto para alguns fatores que permitirão uma melhor compreensão e interpretação desses atributos.

Cabe lembrar que os cálculos para a execução da análise fatorial são complexos e fogem ao objetivo deste trabalho. Na realidade, a utilização de análise fatorial como técnica estatística só se tornou possível depois do advento de computadores e de programas que processam os cálculos de forma rápida e econômica.

Para se executar a análise fatorial deve-se obedecer basicamente a três passos:

2.1 Cálculo das correlações

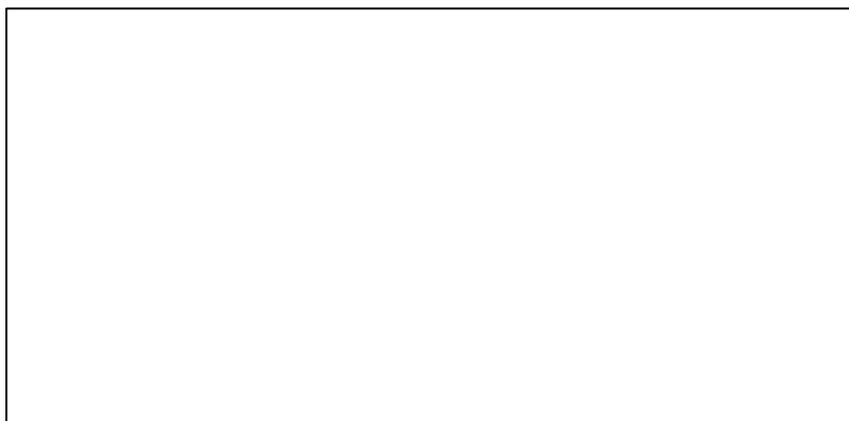
Através da análise de correlação será possível identificar o nível de relacionamento entre as variáveis – duas a duas – conforme metodologia amplamente difundida em livros de econometria básica¹. Para fins de análise cabe esclarecer que o

¹ As calculadoras HP-12C também calculam o coeficiente de correlação diretamente.

coeficiente de correlação entre duas variáveis é um número adimensional que varia de -1 até 1 . Caso o valor do coeficiente de correlação esteja tendendo a um positivo ou negativo – o grau de relacionamento é considerado forte. Aliás, identifica-se como grau forte de relacionamento um coeficiente maior que $0,7$ ou menor que $-0,7$. O sinal positivo indica o tipo de relação entre as variáveis, podendo ser direta (positivo) ou inversa (negativo).

Na Tabela 2 são apresentadas todas as correlações possíveis entre os dados das nove variáveis que se encontram na Tabela 1.

Tabela 2: Correlação linear simples entre as variáveis



As correlações apresentadas na Tabela 2 permitem uma análise parcial dos inter-relacionamentos. Verifica-se, por exemplo, que a variável “Produção total em unidades” mantém relação direta forte com todas as outras variáveis, exceção feita à variável “número de empregados” em que a relação é negativa e, portanto, **inconsistente**. O mesmo fato ocorre com a variável “faturamento total líquido”, o que parece ser razoável admitir já que “produção e faturamento” aparentemente medem a mesma coisa.

Um outro aspecto a ser destacado diz respeito à variável “número de empregados”. Observa-se que a relação inconsistente – negativa – ocorre com todas as outras variáveis inseridas no quadro.

2.2 Extração dos fatores iniciais

Um dos métodos para se extrair os fatores da matriz de correlação é chamado de “método das componentes principais”. Seu objetivo é encontrar um conjunto de

² Em se tratando de produção e empregos, espera-se uma relação direta, ou seja, aumentando-se o nível de produção deve haver uma contrapartida positiva no nível de emprego.

fatores que agreguem linearmente as variáveis na matriz de correlação. Portanto, se as variáveis x_1, x_2, \dots, x_n são altamente correlacionadas entre si, elas serão combinadas para formar um fator.

Por ocasião da extração dos fatores, a análise fatorial fornece uma medida de adequação dos dados (*MSA – Measure of Sampling Adequacy*) para cada uma das variáveis do modelo. Veja as medidas na Tabela 3.

Tabela 3: Medida de adequação dos dados (MSA) para as nove variáveis

--

Um valor de *MSA* inferior a 0,5 indica que a variável em estudo não está adequada à estrutura definida pelas outras variáveis. Neste caso, deve-se optar pela sua eliminação da análise fatorial. No caso em estudo, apenas o *MSA* da variável “exportação em unidades” coincidiu com o limite crítico. Assim sendo, todas as variáveis podem ser utilizadas no modelo em questão.

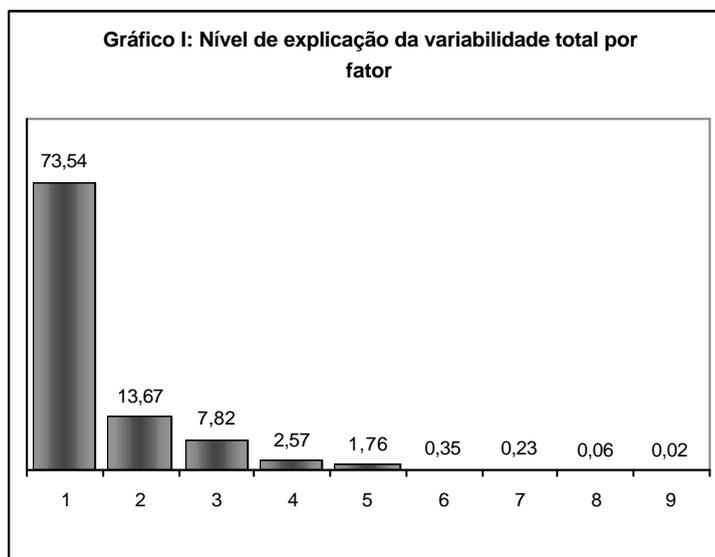
Na seqüência obtiveram-se os valores próprios referentes a cada fator³ e porcentagem da variância explicada para cada um deles.

Tabela 4: Variância total explicada pelos componentes

--

O valor 6,618 apresentado para a componente 1 – denominado *eigenvalue* em inglês – é uma medida de quanto da variância total de todas as medidas realizadas pode ser explicada por aquele fator, ou seja, define a contribuição do fator ao modelo que está sendo construído. Considerando que o usual é utilizar fatores com *eigenvalue* superiores a 1, conclui-se que no exemplo em questão há dois fatores (1 e 2) com valores superiores a 1. Esses dois fatores explicam juntos 87,21% da variabilidade total, conforme se observa no Gráfico I.

³ Na realidade, cada fator pode ser identificado como uma componente principal, pois estamos utilizando o método da componente principal para extrair os “n” fatores.



2.3 Rotação da matriz

Os fatores extraídos ajudam, mas não resolvem completamente o problema de interpretação da análise fatorial. Para melhorar a análise e buscar informações adicionais opta-se pelo giro da matriz original, tecnicamente chamada de *rotação varimax*, cujo objetivo é minimizar o número de variáveis com altas cargas num determinado fator. A tabela 5 apresenta a matriz fatorial depois da rotação.

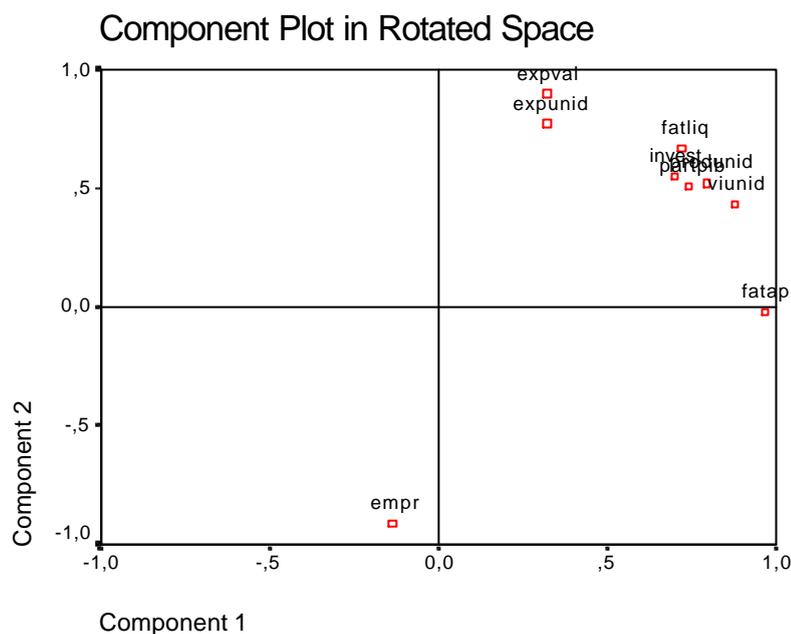
Tabela 5: Composição dos fatores rodados

|--|--|--|--|--|--|--|--|--|--|

Nota-se que em quase todas as variáveis, o fator 1 apresenta carga elevada, portanto, reúne melhor a informação relativa às variáveis. A exceção fica por conta das variáveis exportação em unidades e exportação em valores (com carga baixa), além de número de empregados, que é discriminada por apresentar sinal negativo. No caso do fator 2, a carga é elevada somente para as variáveis exportação em unidades e exportação em valores. Uma melhor visualização pode ser feita através

do Gráfico 2, que apresenta a localização das variáveis num sistema de eixos criado pelos fatores.

Gráfico 2: Relação entre os fatores e as variáveis



Observando o gráfico fica visualmente evidente como as variáveis se agrupam em torno dos fatores 1 e 2. Por outro lado, verifica-se que a variável “número de empregados” não se ajusta a nenhum dos grupos.

Conclusão

Com base no tratamento estatístico efetuado no estudo foi possível verificar que a *análise fatorial* é uma técnica de análise multivariada cujo objetivo principal é representar um conjunto de variáveis originais através de um número reduzido de fatores ou componentes.

No estudo em questão, observou-se que seis das nove variáveis (vendas internas em unidades, produção total em unidades, investimento, faturamento total líquido, faturamento do setor de autopeças e participação no PIB industrial), apresentaram carga elevada em relação ao fator/componente 1. Este agrupamento em

torno de um fator indica que as variáveis estão inter-relacionadas, e podem ser interpretadas como informações específicas de *mercado interno*. Além disso, é provável que algumas delas estejam medindo aproximadamente a mesma coisa, portanto, poderiam ser reduzidas sem comprometer o objeto de estudo.

Para o fator/componente 2, verificou-se que duas entre as nove variáveis (exportação em valores e exportação em unidades), apresentaram cargas elevadas em relação a ele. Neste caso, parece claro que as variáveis devem ser interpretadas como informações específicas de *mercado externo*.

Cabe esclarecer que o objetivo da análise fatorial não é o de isolar as variáveis em grupos distintos. Na prática, a idéia é trabalhar – na medida do possível – com um único fator/componente. Por exemplo, se estamos interessados em montar um modelo de previsão para prever o nível de produção, pode-se utilizar como preditor as vendas externas, em termos de unidades ou de valores, não havendo a necessidade de utilizar as duas variáveis.

Há que se considerar ainda que o fator/componente 1 é responsável por explicar 73,54% da variabilidade total apresentada no modelo em estudo. Ao fator/componente 2 cabe explicar apenas 13,67% da variabilidade total.

Finalmente, verifica-se que a variável “número de empregados” não se ajusta a nenhum dos fatores, mais do que isso, é completamente discriminada, pois apresenta sinal negativo quando deveria apresentar sinal positivo (relação direta). Fica claro, portanto, que o desenvolvimento tecnológico ocorrido no período reduziu consideravelmente as oportunidades de emprego no setor.

Referências Bibliográficas

MAROCO, João. *Análise Estatística com utilização do SPSS*. Lisboa: Silabo, 2003.

MATTAR, Fauze N. *Pesquisa de Marketing*. São Paulo: Atlas, 2001.

PEREIRA, Júlio César Rodrigues. *Análise de dados Qualitativos*. São Paulo: Edusp, 2001.

TRIOLA, Mário F. *Introdução à Estatística*. Rio de Janeiro: LTC Editora, 1999.