

ANÁLISE MULTIVARIADA E O DIAGNÓSTICO DA DINÂMICA DE DESENVOLVIMENTO DOS MUNICÍPIOS DA QUARTA COLÔNIA DO RS (1990-2000)

Valny Giacomelli Sobrinho¹
Alexandre Reis²

RESUMO

Pelo menos desde a década de 1960, economistas e cientistas sociais em geral têm concentrado esforços para encontrar macro-indicadores que lhes informem, com o maior realismo possível, quanto uma cidade, região ou país tem-se desenvolvido. Nesse sentido, dimensões até havia pouco menosprezadas pelos indicadores macroeconômicos tradicionais — como certas variáveis sociais e ambientais — passaram a ser, sobretudo a partir da década de 1980, efetivamente incluídas na medida do desenvolvimento econômico. No entanto, desde que esse conceito mais qualitativo substituiu a noção meramente quantitativa de crescimento econômico, uma série de controvérsias teóricas e práticas teve lugar. Principalmente porque a ponderação das variáveis componentes de qualquer índice de desenvolvimento consiste, invariavelmente, num processo arbitrário, portanto subjetivo. Diversamente, o presente estudo aplica as técnicas de análise multivariada para, ao invés, encontrar, a partir de dados amostrais objetivos, as ponderações reais das variáveis componentes dos índices de desenvolvimento calculados ao longo de uma década (1990-2000) para os municípios da região da Quarta Colônia do estado do Rio Grande do Sul, Brasil. As variáveis utilizadas consistem no valor relativo (percentual) à despesa municipal total por ano de 15 rubricas através das quais essa despesa é alocada. A escolha de itens de despesa e não de receita deve-se a que, no que diz respeito ao orçamento público municipal no Brasil, em particular, as fontes de receita são bastante limitadas (reduzem-se basicamente a dois impostos), ao passo que a aplicação (destinação) dos recursos, que constitui as despesas, informa melhor sobre o perfil do planejamento. Os resultados revelam não só o padrão de desenvolvimento regional em cada período e ao longo da década, como também permitem identificar as variáveis dinâmicas desse processo. A partir desse diagnóstico, então, é possível indicar a direção e o sentido do desenvolvimento local e regional, sem se incorrer no erro comum a muitos planejadores da definição *a priori* dos critérios de ponderação.

Palavras-chave: Indicadores de desenvolvimento; Análise multivariada; Desenvolvimento sustentável; Planejamento regional e urbano.

¹ Professor Assistente do Depto. Ciências Econômicas da UFSM. E-mail: giacomelli@smail.ufsm.br.

² Professor Substituto do Depto. Ciências Econômicas da UFSM. E-mail: alexreisrs@ig.com.br.

1. INTRODUÇÃO

O conceito de sustentabilidade, ou desenvolvimento sustentável, está longe de ser consensual. Como outros conceitos, este também é “parte de um jogo de poder em torno da apropriação do território e de seus recursos, que tem por objetivo legitimar ou deslegitimar discursos e práticas sociais.”(Braga et al., 2003, p. 6). Sua análise e mensuração, portanto, não devem restar dissociadas de uma perspectiva de economia política ecológica, sob pena de se tornarem vazias, estéreis e desprovidas de qualquer significação real.

Desse ponto de vista, o discurso em torno da sustentabilidade reveste-se de duas características fundamentais. Primeiro, trata-a como “sustentabilidade ambiental”, ao pressupor, primordialmente, a sustentabilidade do ambiente biogeofísico. Segundo, impõe a necessidade de uma “sustentabilidade global-local”, como forma de articular os problemas relacionados à sustentabilidade ambiental. Em todo caso, a noção de espaço aqui se encontra fetichizada. Por um lado, o local liga-se abstratamente ao global; por outro, o global é apresentado como um processo econômico anônimo e disperso, que se materializa inexoravelmente. Ulrich Beck suspeita que tal sustentabilidade tenha nascido do desejo desesperado de classes privilegiadas no Ocidente, as quais, tendo alcançado níveis almejados de conforto material, de repente, viram-se, consternadas, em confronto com o risco de, definitivamente, perder-se a imagem utópica que nutriam de uma natureza que nunca existiu (Rangan, 2000).

Seja como for, as narrativas global e local são persuasivas porque apresentam a sustentabilidade como um estado futuro abstraído das condições presentes. Ambas ignoram as regiões e suas histórias geográficas, com o intuito de reforçar sua visão ideal de sustentabilidade. A legitimação do Estado e das políticas públicas galvaniza-se, assim, entre assumir a missão global de salvar o planeta (eco-imperialismo autoritário) e proteger as comunidades locais das forças predatórias da

globalização (eco-paternalismo protecionista). Uma e outra narrativa empregam a linguagem da restrição, da exclusão e do controle autoritário para a futura sobrevivência do planeta ou para o legado lucrativo da terra enquanto propriedade às futuras gerações. Esse tipo de discurso, no entanto, pouco significa para quem mal consegue sobreviver nas circunstâncias presentes e nenhuma propriedade tem a legar às gerações futuras (Rangan, 2000).

Logo, toda medida de sustentabilidade representa, em tese, uma determinada sustentabilidade, politicamente acordada, que pode, inclusive, desde uma perspectiva geográfica, traduzir práticas e concepções fundamentalmente insustentáveis. Afinal, “caracterizar e medir sustentabilidade envolve fazer escolhas sobre como definir e quantificar o que está sendo desenvolvido, o que está sendo sustentado e por quanto tempo. Os objetivos, indicadores e metas de sustentabilidade, são derivados dessas escolhas.” (Parris & Kates, 2003, p. 572)³. *Objetivos* são enunciados qualitativos específicos mas de conteúdo amplo, derivados de duas vertentes básicas, tais como indicam as colunas do Quadro 1, abaixo. *Indicadores* são medidas quantitativas escolhidas para avaliar o progresso em direção a ou para longe de um determinado objetivo. *Metas* empregam indicadores para delimitar objetivos no tempo e no espaço, definindo-lhes cronogramas de execução. Mudanças nos indicadores ao longo do tempo descrevem *tendências*, que são determinadas por *forças motrizes (driving forces)* e *reações planejadas (policy responses)* e, por conseguinte, influenciam a capacidade de cumprimento das metas (Parris & Kates, 2003).

³ Characterizing and measuring sustainability involves making choices about how to define and quantify what is being developed, what is being sustained, and for how long. The goals, indicators and targets of sustainability ... are derivative of these choices.

Por isso há que se distinguir entre indicadores, forças motrizes e respostas do planejamento. Nem que seja apenas para esclarecer por que um indicador, tomado isoladamente, é, em si, pouco significativo e, mesmo, ambíguo. De modo que, por esse motivo, todo diagnóstico e definição de prioridades deveriam assentar-se numa bateria de indicadores, que permitisse vislumbrar diversas orientações. Reagrupados, esses diversos indicadores podem ser associados a um dado universo, tomando-se cada variável independentemente das outras, e construindo-se, assim, um perfil ou balanço a partir de quadros estatísticos representando as médias, as distribuições, as tendências, abertas à análise (Bartoli, 1996).

Em essência, qualquer que seja a medida que se utilize, a sustentabilidade pressupõe a questão geográfica de acesso aos recursos regionais e ao bem-estar social no “presente contínuo”, mais do que em algum passado idílico ou futuro distante e utópico (Rangan, 2000). Por isso o que se indica como sustentável em determinada região não necessariamente o é em outra. Nesse sentido, a definição de políticas públicas voltadas para o desenvolvimento regional depende do maior conhecimento das especificidades das regiões e das delimitações precisas do espaço a ser abordado. Por exemplo, em termos das oportunidades de investimento, das restrições ambientais e dos problemas sociais. Tal regionalização deveria considerar as desigualdades não apenas entre macrorregiões mas também entre ecossistemas (Siqueira & Siffert Filho, 2001).

Este estudo constitui um esforço de aplicação do que Rangan (2000, p. 123) considera o propósito fundamental da ecologia política: “explicar *os resultados biogeográficos das relações sociais* em um contexto de configurações políticas e espaciais diferenciadas.” Não se trata de definir a “norma” em matéria de condições

de sustentabilidade. Ao contrário, trata-se de reconhecer a complexidade em estabelecê-la — o que, entretanto, é muito diverso de renunciar a ela.

IncurSIONAR pela complexidade sem enovelar-se nela começa por assentar a definição de metas em bases menos subjetivas. Em geral, metas emergem de escolhas feitas com base no consenso ou na negociação. Nesse caso, envolvem muitos interesses e aspectos normativos que podem comprometer seu cumprimento ou sustentabilidade. Noutras palavras, elas até podem desfrutar de legitimidade (atenção à diversidade de crenças, valores e interesses das partes envolvidas), mas não necessariamente de “saliência” (medidas relevantes à formulação de políticas públicas) ou credibilidade (consistência técnica e científica). Desse modo, arriscam-se a frustrar o alcance dos objetivos primários (Quadro 1), para os quais foram estabelecidas. Por isso, cada vez mais, as metas devem deslocar-se no sentido da teoria científica e da pesquisa (Parris & Kates, 2003).

Quadro 1 – Taxonomia dos objetivos do desenvolvimento sustentável

O que deve ser sustentado	O que deve ser desenvolvido
Natureza Terra Biodiversidade Ecossistemas	População Sobrevivência infantil Expectativa de vida Educação Equidade Igualdade de oportunidades
Suporte da vida Serviços dos ecossistemas Recursos Meio ambiente	Economia Riqueza Setores produtivos Consumo
Comunidade Cultura Grupos Lugares	Sociedade Instituições Capital social Estados Regiões

Fonte: Parris & Kates, 2003, p. 561

A substituição da arbitrariedade na seleção e ponderação dos critérios que medem o desenvolvimento regional pela fidelidade aos dados objetivos é precisamente o que se pode conseguir a partir da aplicação das técnicas de análise multivariada. Não há critérios preestabelecidos. São os dados que, segundo uma escala de variâncias máximas, revelarão os critérios a serem considerados.

O que se vai fazer aqui é aplicar cada uma dessas técnicas à análise do processo de desenvolvimento regional dos municípios da Quarta Colônia, na região Central do estado (RS), no período de 1990 a 2000. O objetivo é chegar a um diagnóstico dos elementos (variáveis) dinâmicos(as) desse processo. Para tanto, será calculado um índice de desenvolvimento para cada município. Esse valor informa, vetorialmente (i. e., em sentido positivo ou negativo), o desempenho municipal no período. A *análise de componentes principais* revela — a partir do percentual da variância total explicada que cada fator representa — as ponderações (*factor loadings*) que cada um deles efetivamente carrega na composição do índice. O que se discutirá, então, mais adiante, é quantas dimensões (fatores) devem integrar o indicador. De um lado, muitas dimensões (fatores) dificultam sua interpretação; de outro, poucas dimensões (fatores) podem não ser suficientes para representar uma porção significativa da variância total explicada (Kubrusly, 2001).

2. METODOLOGIA

Todo tratamento estatístico paramétrico lida com variabilidades em torno de um valor médio. Com base na explicação dessa variabilidade, é que é possível inferir (ou predizer), com relativa certeza, o comportamento futuro de um conjunto de variáveis ou generalizar, a partir de casos particulares (amostras), o comportamento presente desse mesmo conjunto. Existem variáveis (dependentes), no entanto, que

são compostas por uma série extensa de outras (independentes). Esse é o caso da variável “desenvolvimento” (econômico, socioeconômico, ambiental, sustentável etc.). Quanto mais acurada se pretende essa medida, mais complexa se torna sua análise. Nessas circunstâncias, a redução da complexidade passa pela redução dos dados (variáveis independentes) a serem analisados. O objetivo é descobrir ou detectar a estrutura subjacente (ou latente) às relações entre as variáveis. Para tanto, removem-se as variáveis julgadas redundantes, devido à alta correlação que mantêm entre si. Como são altamente correlacionadas, variam com amplitudes muito semelhantes (variâncias ou desvios-padrão semelhantes). Assim, podem-se fundi-las em uma só, proporcionando, dessa forma, a redução desejada no número de variáveis (independentes) a ser analisado.

A análise de componentes principais e a análise fatorial são técnicas da análise estatística multivariada que consistem em agrupar, hierarquicamente, em novas variáveis (independentes) denominadas “componentes” ou “fatores”, as variâncias amostrais de cada variável independente. Na análise de componentes principais, aplicam-se os mesmos métodos utilizados na análise fatorial tradicional — também conhecida como “análise de fatores comuns”. A diferença é que, na primeira, analisa-se a variância *total*; na última, a variância *comum* (ou *comunalidades*). Neste caso, busca-se construir uma matriz de correlação, contendo, na diagonal principal, as *comunalidades estimadas*. Quanto mais próximas de 1, maior é “a proporção da variância de cada variável que pode ser explicada (reunida) pelos (em) fatores comuns” (Mardia et al., 1979, p. 263). Já na análise de componentes principais, trabalha-se com uma matriz de correlação cuja diagonal principal é unitária (UCLA, 2004).

Basicamente, a análise de componentes principais empenha-se em obter uma “sumarização parcimoniosa” dos dados. Para tanto, procura encontrar a “combinação linear padronizada” (ponderação) das variáveis originais que maximize a variância total (Mardia et al., 1979). Cada componente (autovetor ou *eigenvector*) carrega uma fração (autovalor ou *eigenvalue*) da variância total correspondente à sua respectiva ponderação. Da mesma forma, cada variável, dentro da componente principal (autovetor), recebe uma ponderação (coeficiente fatorial) equivalente a seu respectivo comportamento amostral. Logo nem todas as variáveis exercerão a mesma influência sobre uma determinada componente, uma vez que nem todas variarão com a mesma intensidade através das amostras. Desse modo, é possível identificar quais variáveis exercem maior influência sobre cada componente. Além do mais, como se mencionou antes, a combinação linear (autovalores) encontrada para as componentes contempla necessariamente a maior parcela da variância total. Assim, combinadas dessa forma, as componentes principais podem ser utilizadas para construir índices que resumem os dados através de algumas poucas dimensões (componentes ou fatores). Essa combinação ótima (no sentido da variância máxima) garante, indiscutivelmente, maior consistência com a realidade do que qualquer outro sistema arbitrário de ponderação, tal como média aritmética simples.

Similarmente, a análise fatorial consiste num “modelo matemático que tenta explicar a correlação entre um vasto conjunto de variáveis em termos de um pequeno número *fatores* subjacentes.” (Mardia et al., 1979, p. 255). Esses fatores, no entanto, não são observáveis diretamente. Tal pressuposto é especialmente útil a objetos de estudo — como desenvolvimento econômico, inteligência etc. — cuja mensuração exata e definição conceitual são difíceis. Não obstante a identificação da maior variabilidade permite “separar” melhor as amostras (municípios), evidenciando,

dessa maneira, as diferenças entre elas (eles) e possibilitando inferências mais refinadas a seu respeito.

Para isso, a variância total é decomposta (ou fatorada) em diversos elementos os quais estão indicados no Quadro 2.

Quadro 2 – Decomposição ou fatoração da variância total na análise fatorial

Variância total	$h_i^2 + b_i^2 + e_i^2$
Confiabilidade	$h_i^2 + b_i^2$
Comunalidade	$h_i^2 = \sum \lambda_{ij}^2$
Variância única (<i>uniqueness</i>)	$u_i^2 = b_i^2 + e_i^2$
Especificidade	b_i^2
Erro	e_i^2

Fonte: UCLA, 2004

“A comunalidade representa a variância que a variável X_i compartilha com as outras variáveis, via fatores comuns.” (Mardia et al., 1979, p. 257). Isto é, a covariância entre a variável X_i e o fator F_j é λ_{ij}^2 . Matematicamente:

$$\text{Cov}(X_i, F_j) = \lambda_{ij}^2$$

Os fatores comuns são padronizados para que cada um tenha variância igual à unidade. Assim:

$$\sigma_{ii} = \sum_{j=1}^k \lambda_{ij}^2 + \psi_{ii}$$

$$1 = h_i^2 + \psi_{ii}$$

Portanto, quanto mais próximo h_i^2 estiver de 1, menor é a variância única ou específica ψ_{ii} . Do contrário, ψ_{ii} “explica a variabilidade em X_i não compartilhada com as outras variáveis” (Mardia et al., 1979, p. 257). O “modelo fatorial” pode ser então descrito como:

$$\begin{aligned} X_1 - \mu_1 &= (\lambda_{11}F_1 + \lambda_{12}F_2 + \dots + \lambda_{1k}F_k) + u_1 \\ X_2 - \mu_2 &= (\lambda_{21}F_1 + \lambda_{22}F_2 + \dots + \lambda_{2k}F_k) + u_2 \\ &\vdots \\ X_k - \mu_k &= (\lambda_{k1}F_1 + \lambda_{k2}F_2 + \dots + \lambda_{kk}F_k) + u_k \end{aligned}$$

Equação 1

Ou, vetorialmente:

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix}_{k \times 1} - \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{bmatrix}_{k \times 1} = \underbrace{\begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1k} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{k1} & \lambda_{k2} & \dots & \lambda_{kk} \end{bmatrix}}_{k \times k} \underbrace{\begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_k \end{bmatrix}}_{k \times 1} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}_{k \times 1},$$

Equação 2

que, matricialmente, pode ser abreviado por:

$$\mathbf{X} - \boldsymbol{\mu} = \boldsymbol{\Lambda} \mathbf{F} + \mathbf{u}$$

Equação 3

Na Equação 2, à semelhança do que ocorre com outros tipos de médias ponderadas ou com a média aritmética simples, cada linha da matriz de carregamentos $\boldsymbol{\Lambda}$ deve totalizar 1 — que é a variância padronizada dos fatores comuns. Ou seja:

$$\sum_{j=1}^k \lambda_{ij}^2 = 1$$

Equação 4

A padronização do modelo descrito pela Equação 3 também requer que o vetor de médias amostrais $\boldsymbol{\mu}$ seja nulo, isto é:

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Como a variância total $[(\mathbf{X} - \boldsymbol{\mu})^2]$ ou desvio-padrão $[(\mathbf{X} - \boldsymbol{\mu})]$ associados a cada variável devem ser iguais à unidade, e o erro \mathbf{u} , associado às estimativas da influência dos fatores não-observáveis sobre essa variância é estatisticamente desprezível, o modelo da Equação 3 reduz-se a:

$$\mathbf{1} = \boldsymbol{\Lambda} \mathbf{F}$$

Logo:

$$\mathbf{F} = \boldsymbol{\Lambda}^{-1}$$

Equação 5

Todavia os valores de \mathbf{X} são fornecidos pelas amostras, a partir dos dados. O que o modelo fatorial quer descobrir, no entanto, são os elementos “ocultos” dentro do vetor \mathbf{F} . Como se viu, eles não são diretamente observáveis a partir dos dados, mas podem ser estimados com uma margem estatisticamente desprezível de erro (\mathbf{u}). Assim, a Equação 3, com o vetor de médias amostrais $\boldsymbol{\mu}$ nulo, é mais apropriadamente especificada como:

$$\mathbf{F}_{(k \times 1)} = \boldsymbol{\Lambda}^{-1}_{(k \times k)} \mathbf{X}_{(k \times 1)}$$

Equação 6

Com $\boldsymbol{\mu}$ nulo, \mathbf{X} é que determina o tamanho da variância. Assim, há que se encontrar a combinação linear normalizada que, ao mesmo tempo, maximize a variância de \mathbf{F} e obedeça à restrição estabelecida pela Equação 4. Trata-se, portanto, de um problema de otimização, que pode ser descrito como⁴:

$$\text{Max} Z = E(\boldsymbol{\Lambda}^{-1} X)^2 = \text{Var}(\boldsymbol{\Lambda}^{-1} X) = (\boldsymbol{\Lambda}^{-1})^2 \text{Var}(X) = (\boldsymbol{\Lambda}^{-1})' \boldsymbol{\Sigma} (\boldsymbol{\Lambda}^{-1})$$

sujeito a⁵:

$$(\boldsymbol{\Lambda}^{-1})' (\boldsymbol{\Lambda}^{-1}) = \mathbf{1}$$

Inserindo-se a restrição na função-objetivo Z , através do multiplicador de Lagrange λ , tem-se:

$$Z = (\boldsymbol{\Lambda}^{-1})' \boldsymbol{\Sigma} (\boldsymbol{\Lambda}^{-1}) - \lambda [(\boldsymbol{\Lambda}^{-1})' (\boldsymbol{\Lambda}^{-1}) - 1]$$

Equação 7

Maximiza-se Z em relação a $\boldsymbol{\Lambda}^{-1}$, através da condição de primeira ordem. Isto é:

$$\begin{aligned} \frac{\partial Z}{\partial \boldsymbol{\Lambda}^{-1}} &= 0 \\ \frac{\partial Z}{\partial \boldsymbol{\Lambda}^{-1}} &= 2\boldsymbol{\Sigma} \boldsymbol{\Lambda}^{-1} - 2\lambda \boldsymbol{\Lambda}^{-1} = 0 \\ \boldsymbol{\Sigma} \boldsymbol{\Lambda}^{-1} - \lambda \boldsymbol{\Lambda}^{-1} &= 0 \\ (\boldsymbol{\Sigma} - \lambda \mathbf{I}) (\boldsymbol{\Lambda}^{-1}) &= 0 \end{aligned}$$

⁴ Conforme T. W. Anderson (1984), *An introduction to multivariate statistical analysis*, p. 453.

⁵ Equivalente matricial da Equação 4.

Equação 8

Se o determinante de uma matriz qualquer é nulo, as linhas e/ou colunas dessa matriz *não* são linearmente independentes. Isso significa que há, na verdade, mais incógnitas (colunas) que equações (linhas) e que, portanto, o sistema não tem solução ou é indeterminado. Por outro lado, se qualquer das linhas ou colunas de uma matriz, ao ser decomposta em vetores (linha ou coluna), puder ser expressa como uma combinação linear de pelo menos um outro desses vetores (linha ou coluna), isso também é uma outra maneira de dizer que o sistema apresenta mais incógnitas que equações e que, outra vez, não terá solução ou será indeterminado. Algebricamente, a independência linear — por conseguinte a solução do sistema — requer que:

$$[\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_k] \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_k \end{bmatrix} = 0$$

$$\lambda_1 F_1 + \lambda_2 F_2 + \dots + \lambda_k F_k = 0$$

$$\sum_{i,j=1}^k \lambda_{ij} F_j = 0$$

Equação 9

Logo, se o determinante da matriz que pré-multiplica Λ^{-1} na Equação 8 for nulo, isso assegura que os coeficientes fatoriais encontrados pela matriz inversa de carregamentos Λ^{-1} não representarão combinações lineares entre dois ou mais fatores, qualquer que seja o vetor \mathbf{X} na Equação 6.

Então, como antes, o primeiro fator vai reunir as maiores variabilidades amostrais para cada variável (independente); o segundo fator, as imediatamente inferiores; e assim sucessivamente. Ao mesmo tempo, a decomposição (vertical, através das linhas) de cada variável dentro de cada fator (coeficientes fatoriais ou

coeficientes da componente principal) permite verificar quais delas o afetam mais. O somatório (vertical) dos produtos desses pesos relativos das variáveis dentro de cada fator (coeficientes fatoriais) ponderados pela recíproca da raiz quadrada do autovalor associado àquele fator (Kubrusly, 2001) e pelo valor original da variável em cada elemento da amostra (objeto) resultará num escore fatorial para cada um desses elementos — no presente estudo, municípios (SPSS for Windows, 2001). Ou, ainda, “os carregamentos fatoriais [ou coeficientes fatoriais] estimados são os coeficientes da componente principal amostral (autovetores de \mathbf{R}^6) escalonados pela raiz quadrada dos correspondentes autovalores.” (Johnson & Wichern, 1982, p. 390).

Seja como for, os escores fatoriais respondem também, eles próprios, pela fração da variância total (autovalor) que compete àquele dado fator. Com efeito, essa conclusão já estava presente na Equação 9. Ali, fica demonstrado que a própria construção dos fatores é impossível sem o conhecimento dos autovalores, extraídos a partir da Equação 8.

A distribuição da variabilidade é tal que os fatores iniciais carregam as maiores frações da variância total explicada. Por essa razão, é possível reduzir-se uma infinidade de variáveis explicativas (independentes) a uns poucos fatores (novas variáveis explicativas), visto que a fração da variância total desprezada é muito pequena; logo estatisticamente insignificante (Tabela 1).

Como já se disse, tal economia é útil à análise de variáveis dependentes complexas — como o desenvolvimento econômico —, que envolvem variáveis independentes qualitativas e/ou quantitativas. Os fatores agrupam as variáveis independentes com base em seus elementos comuns (comunalidades). Assim, é possível identificar qual(is) dos critérios (fatores) utilizados mais marcadamente

influencia(m) o desenvolvimento econômico de cada município de uma determinada região.

Através da análise de componentes principais, descrita por Johnson e Wichern (1982), pode-se, sem perda de informação (variáveis independentes qualitativas ou fatores), reduzir a matriz de fatores original a algumas poucas componentes (colunas) que contemplam as maiores variâncias. Como cada linha dessa matriz vai representar um município do universo estudado, o valor assumido pela componente correspondente (i. e., através das colunas) permite identificar qual(is) o(s) critério(s) [fator(es)] que mais ou menos peso deposita(m) no processo de desenvolvimento municipal. Quanto mais elevado, em módulo, for o valor assumido pela variável dentro da componente principal, maior é o efeito que aquela exerce nesta. A mesma comparação pode ser feita entre as linhas. Se, aqui também, cada linha representa uma variável independente, a análise, agora, pode ser estendida para identificar aquelas que mais influência exercem sobre cada critério (fator). Desse modo, o peso de cada fator no desenvolvimento municipal/regional pode ser rastreado ao nível das variáveis (independentes) mais influentes. É precisamente isso que possibilita a formulação de políticas de desenvolvimento locais/regionais.

Tabela 1 – Síntese dos procedimentos da análise fatorial

Variáveis $i = 1, \dots, k$	Comunalidades (λ^2_{ij})			Fatores (F_j)		
	$j = 1$...	k	$j = 1$...	k
	λ^2_{i1}	...	λ^2_{ik}	F_1	...	F_k
X_l	λ^2_{l1}	...	λ^2_{lk}	Max $\{\lambda^2_{lk}\}$...	Min $\{\lambda^2_{lk}\}$
\vdots	\vdots		\vdots	\vdots		\vdots
X_k	λ^2_{k1}	...	λ^2_{kk}	Max $\{\lambda^2_{kk}\}$...	Min $\{\lambda^2_{kk}\}$
Z^*				Max $\sum_{i=1}^k \sum_{j=1}^k \lambda^2_{ik}$...	Min $\sum_{i=1}^k \sum_{j=1}^k \lambda^2_{ik}$

(*) Função-objetivo

3.FONTE DE DADOS

Os dados analisados referem-se às razões, para cada município, entre as despesas correspondentes a 15 diferentes rubricas e a despesa municipal total no período. A fonte dos dados é o Tribunal de Contas do Estado do RS. O emprego das razões entre itens de despesa e despesa total num mesmo período torna-se simultaneamente vantajoso e eficiente, porquanto oferece uma idéia clara da participação relativa de cada item de despesa no total da despesa, sem a necessidade de recorrer ao uso de deflatores, que tornariam o cálculo mais trabalhoso.

O motivo para se haverem escolhido itens da despesa e não da receita é que a despesa reflete melhor as preocupações do gestor. No que se refere ao orçamento público, em particular, a aplicação (destinação) dos recursos (despesa) informa mais a respeito da capacidade de planejamento do que a origem deles (receita), cujas fontes, no caso municipal, são bastante limitadas (IPTU e ISSQN fundamentalmente). A despeito de certas vinculações legais, a alocação desses recursos entre fins alternativos (despesas) é que reflete, com mais propriedade, o que é ou não prioritário do ponto de vista dos planejadores públicos.

4.ANÁLISE E DISCUSSÃO DOS RESULTADOS

O índice 1990-2000 de desenvolvimento para cada município da região foi construído da seguinte maneira:

- a) Calculando-se, para cada ano, as variâncias de cada variável (item de despesa) através das amostras (municípios). A partir daí, chegou-se a um índice de desenvolvimento anual para cada município, por meio de análise fatorial, em que os fatores apontam o peso relativo de cada item de despesa

(variáveis independentes) no desempenho (índice) municipal (variável dependente) naquele ano específico. Assim, para cada ano, gerou-se uma tabela reunindo os índices de cada município, num total de 11 tabelas⁷. Em cada uma delas, o número de fatores compondo o índice anual municipal varia, assim como o peso (coeficientes fatoriais) das variáveis na composição de cada fator. Por si só, isso atesta a ausência de arbitrariedade nas ponderações, bem como a fidelidade da técnica aos dados objetivos. Em cada ano, o índice municipal corresponde à soma dos escores fatoriais

$$\hat{f}_j = (\Lambda' \Lambda)^{-1} \Lambda' (X_j - \bar{X})$$

Equação 10

obtidos para cada município. Para dados padronizados (média = 0 e variância = 1) divide-se a Equação 10 por $\sqrt{s_{ii}}$.

- b) A seguir, transformaram-se os anos em variáveis cujo valor passou a ser, agora, o índice de desenvolvimento encontrado para o município correspondente (Tabela 2). As variâncias, portanto, referem-se, aqui, ao comportamento do índice municipal ao longo da década. A análise fatorial indica, agora, os anos que mais contribuíram para o desempenho municipal, com os fatores representando o peso relativo de cada ano no somatório dos índices municipais anuais. Assim, gerou-se uma única tabela, sintetizando num índice calculado para cada município o seu desempenho ponderado ao longo da década.
- c) Confrontando-se a Tabela 2 com a **Tabela 3**, é possível rastrear, através da série amostral (municipal) de índices anuais, as causas determinantes do desempenho de cada município e da região como um todo ao longo do

⁷ Visto que se incluíram nesses cálculos tanto o ano de 1990 quanto o de 2000.

período. Noutras palavras, mesmo que o município tenha tido bom desempenho (índice mais elevado) em anos pouco favoráveis ao desempenho de todo o período, com certeza, ele não figurará entre os mais bem classificados pelo índice global (1990-2000). Esse rastreamento pode ser feito com base na Tabela 4.

Tabela 2 – Índices de desenvolvimento municipais anuais (1990-2000)

Municípios	Índices ¹										
	1990	1991	1992	1993 ³	1994	1995	1996	1997	1998	1999	2000
1) Arroio Grande	-2,840880	0,873870	1,61760	1,79559	0,31448	-2,036050	1,244080	1,105210	-1,287170	-2,404980	
2) Dona Francisca	0,506540	-1,613700	-2,23163	-1,05849	-1,64889	-2,254010	0,456290	0,181480	1,283530	2,069220	0,468040
3) Faxinal do Soturno	1,301270	2,222320	-0,03687	0,57666	0,13861	-0,706600	-2,789430	2,964070	3,817460	-0,411330	-2,630560
4) Ivorá	0,654570	-1,537950	0,95981	0,60412	-4,97657	-2,066700	-2,008800	1,983440	-1,009900	-1,526110	-0,117970
5) Nova Palma	1,995790	1,113000	1,22992	2,01432	-0,61751	-0,623210	-0,828070	-5,650190	-3,498300	-0,200630	2,456790
6) Pinhal Grande				-4,82495	1,69247	1,122190	0,504860	-2,082970	0,367830	-1,858660	-3,914810
7) Restinga Seca	-2,840880	-3,489370	-3,29764	-1,91669	1,84495	1,139800	-2,639000	1,506700	0,197700	3,761320	1,933370
8) São João Polésine				1,41664	2,10466	2,297790	2,746300	0,193240	-2,634830	-0,455210	0,010690
9) Silveira Martins	1,223610	2,431830	1,75880	1,39278	1,14779	3,126760	3,313740	-0,947430	2,592040	3,164590	2,649010
10) Itaara ²								0,746490	0,171630	-2,138200	-0,854620

(1) As células em branco indicam “missing values” (valores ausentes). Na análise fatorial, foram excluídas conforme o procedimento “listwise”.

(2) Município inexistente ou cujos dados ainda não estavam disponíveis até 1996.

Tabela 3 – Índices globais municipais da Quarta Colônia (RS) durante o período de 1990-2000

Períodos (anos)	F ₁	F ₂	F ₃	Total
	Início	Meio-fim	3/4	
T1990	0,227563	-0,05459	0,025837	
T1991	0,251105	0,102326	0,159761	
T1992	0,224739	-0,00674	-0,01436	
T1993	0,222628	-0,01191	-0,05577	
T1994	-0,01396	0,262565	0,067562	
T1995	0,050166	0,283328	0,030456	
T1996	0,111552	0,192491	-0,04995	
T1997	-0,05145	-0,02268	0,338213	
T1998	0,077842	0,16015	0,412384	
T1999	-0,10464	0,254256	-0,02011	
T2000	-0,05468	0,135509	-0,34177	
Autovalores (λ)	4,106579	3,28314	2,377303	9,767022
% Variância	37,33254	29,84672	21,61184	88,79111

(*) Municípios cujos índices estão ausentes da Tabela 2 para alguns anos.

TABELA 4 – Causalidade temporal nos fatores*

MUNICÍPIOS	Escores fatoriais			ÍNDICE
	$\wedge f_1$	$\wedge f_2$	$\wedge f_3$	Ranking
Silveira Martins	0,95217	1,56091	0,01963	2,532710
Faxinal do Soturno	0,53086	-0,31203	1,64805	1,866880
Arroio Grande*	—	—	—	—
Pinhal Grande*	—	—	—	—
São João Polêsine*	—	—	—	—
Itaara*	—	—	—	—
Dona Francisca	-0,56671	-0,20006	0,0634	-0,703370
Nova Palma	0,7395	-0,321	-1,50017	-1,081670
Restinga Seca	-1,71466	0,65055	-0,1496	-1,213710
Ivorá	0,05885	-1,37837	-0,0813	-1,400820
Quarta Colônia				0,000020

(*)Component score coefficient matrix (matriz dos coeficientes de escore fatorial, rotacionada pelo método Varimax)

Primeiro, os três fatores explicam perto de 90% da variância amostral total. Os valores em negrito identificam as variáveis (anos) com os maiores carregamentos dentro e através dos fatores. Isto é, aquelas que mais influência exercem sobre o fator e sobre a variância amostral total, respectivamente. São precisamente esses carregamentos que permitem denominar os fatores e diagnosticar os desenvolvimentos municipais e regional no período.

Note-se que, na Quarta Colônia como um todo, ao longo do período, os anos do início da década foram decisivos. São eles que carregam a maior parcela da

variância (37%). Quer dizer, os municípios que conquistaram índices elevados aí, devem alcançar altos índices globais. Não com tanta intensidade, mas a metade e o final da década também podem determinar, à razão de cerca de 30% da variabilidade total, um bom desempenho global. Por fim, com pouco mais de 20% da carga de variabilidade, o terceiro quartel da década participa mais fracamente do processo de desenvolvimento municipal no período.

Com a ajuda da **Tabela 3**, da Tabela 4 e da **Tabela 2**, nessa ordem, é possível chegar-se rapidamente a um diagnóstico de razoável consistência. Por exemplo: a **Tabela 3** informa que Silveira Martins alcançou o maior índice da região no período 1990-2000 (2,532710) e que, ainda, obteve seu primeiro maior escore no fator 2 (1,56091) seguido pelo do fator 1 (0,95217). Agora, então, com base na Tabela 4, pode-se inferir que Silveira Martins tenha conquistado seus maiores índices de desenvolvimento na metade e no final da década (1994-1996, 1999-2000) e índices moderados no início da década (1990-1993). A confirmação dessas hipóteses deve ser finalmente buscada na **Tabela 2**. De fato, elas se revelam bastante consistente com os dados.

O mesmo algoritmo pode ser aplicado aos demais municípios e à região como um todo. Como se vê na **Tabela 3**, a Quarta Colônia manteve-se estagnada no período (0,000020). A maioria de seus municípios acusaram desempenhos negativos, à exceção de Silveira Martins e Faxinal do Soturno, que evitaram o declínio regional. Ao contrário de Silveira Martins, contudo, Faxinal do Soturno pontuou mal no fator 2 (metade e final da década). Obteve escore negativo onde Silveira Martins pontuou melhor. Similarmente, no início da década, Silveira Martins teve um desempenho (escore) quase duas vezes melhor que Faxinal do Soturno. Ao longo do período, Faxinal do Soturno conseguiu, no terceiro quartel da década, o maior escore fatorial de toda a região em qualquer período. Apesar disso, a “escolha” recaiu sobre a época errada, uma vez que o fator 3 é justamente o que responde pela menor parcela da variância total explicada. Logo seu poder de propulsão ou de implosão é bem mais baixo. Com certeza, a dinâmica do desenvolvimento regional ao longo da última década não pode ser atribuída a esse período (terceiro quartel).

Um passo adiante no diagnóstico pode ser dado retornando-se à origem. Quer dizer, procurando-se identificar que itens de despesa, afinal, foram decisivos para o desempenho municipal (ou regional) em determinado período. Não cabe aqui exibir

as 11 matrizes de ordem (10 x 15) com os itens de despesa relativos de cada município. Contudo, só para tomar um exemplo, convém, primeiro, examinar o escore mais elevado (f_2) de Silveira Martins. Como esse escore deposita seu peso nos anos da metade (1994, 1995 e 1996) e do final da década (1999 e 2000), deve-se buscar nesses períodos quais os itens de despesa que, de acordo com a composição fatorial, mais contribuíram para o valor do índice de desenvolvimento anual.

Em 1994, por exemplo, o único escore positivo de Silveira Martins foi na componente (fator) $F_2 = 2,49948$. A Tabela 5, abaixo, equivalente, para o ano de 1994, à TABELA 4, permite identificar a razão (em termos de itens de despesa agora) desse bom desempenho.

Tabela 5 – Causalidade do dispêndio para o ano de 1994*

Variáveis	Fatores ou componentes principais				
	F_1	F_2	F_3	F_4	F_5
LEGIS	-0,0001	0,023405	-0,20673	-0,27399	-0,23483
JUD	0,050547	0,213803	-0,00686	0,076557	0,238297
ADM	-0,11465	0,367918	-0,00784	0,01471	-0,01312
AGRIC	-0,14172	0,044568	-0,21959	0,361205	-0,08319
COMUNIC	0,192636	0,080089	0,158463	-0,00654	-0,05534
SEGPUB	0,271338	-0,07048	-0,09141	-0,02325	0,208076
DESENV	0,305492	-0,01293	-0,09265	-0,03246	-0,01871
EDUC	-0,03313	-0,08174	-0,05479	-0,01789	0,462211
ENEREC	-0,06537	-0,11824	-0,02826	-0,37495	-0,01838
HAB	0,011354	-0,04706	0,365434	0,007663	-0,07006
INDCOM	-0,04306	-0,20955	-0,02746	0,306108	-0,05084
SAÚDE	0,042298	0,34934	0,039225	0,082984	-0,16758
TRAB	0,269731	-0,07155	-0,02443	-0,00455	-0,24818
TRANSP	-0,09397	0,082214	0,420654	-0,05473	-0,02505

(*) Component score coefficient matrix (rotacionada pelo método Varimax)

As hipóteses que se levantaram antes, na análise da causalidade temporal, podem ser formuladas de maneira semelhante aqui também. Na certa, o alto escore de Silveira Martins em 1994 se deveu a elevados gastos com administração (ADM) e saúde — as variáveis com maiores coeficientes fatoriais (absolutos) em F_2 . A esta altura, o planejamento já pode avançar no sentido de elaboração de prognósticos e cenários de desenvolvimento regional e municipal.

5. CONCLUSÃO

Possivelmente, a contribuição mais expressiva desta pesquisa diz respeito à reflexão acerca da natureza metodológica dos índices de desenvolvimento. Usualmente, esses índices são concebidos para medir um “ideal” de desenvolvimento. Ou melhor, a distância ou proximidade a que um país ou região se encontra desse “ideal”. Um problema prático que imediatamente emerge daí se refere às bases sobre as quais esse “ideal” é erigido, porquanto um “ideal” sempre comporta elementos subjetivos. Dessa perspectiva, constitui-se, pois, num conceito relativístico, uma vez que o que pode parecer ideal a uns não necessariamente o será para outros.

Seja como for, mesmo de uma perspectiva mais teórica, como na do conceito de “tipo ideal” de Max Weber, resultam complicações empíricas. O “tipo ideal” de Weber não descreve um curso concreto de ação, mas um desenvolvimento normativamente ideal. “É um conceito vazio de conteúdo real: ele depura as propriedades dos fenômenos reais, desencarnando-os pela análise, para depois reconstruí-los.” (Tragtenberg in Weber, 1997, p. 8). Serve como “pauta de contrastação”. Não constitui hipótese nem proposição, não podendo, portanto, ser falso nem verdadeiro, mas válido ou não-válido. Somente os resultados que com eles se obtenham na análise da realidade social podem dar a medida de sua conveniência, pois trata-se de tipos puramente conceituais, jamais se encontrando na realidade em toda a sua pureza (Tragtenberg in Weber, 1997). Logo, admitidamente, jamais haverá condições objetivas para se atingi-los. Em resumo, o ideal definido nesses termos não é necessariamente o possível. Como tal, pouco serve aos propósitos do planejamento, porquanto o planejamento se trata de uma atividade aplicada, com objetivos traduzíveis em metas factíveis, mais que teóricas.

De resto, a técnica empregada aqui requer um esforço continuado, conquanto toda técnica deve ser escrava dos dados; não o contrário. Seja como for, a análise fatorial é, em princípio, empregada em caráter exploratório, para identificar estruturas invariantes não-observáveis. Presta-se, sobretudo, à compreensão de variáveis categóricas ou dicotômicas. Por isso, a rigor, idealmente, procura por padrões dicotômicos dispersos ou ocultos no amplo espectro do universo multivariado.

É útil nos casos em que se verifica a pressuposta normalidade dos dados, com variáveis simetricamente distribuídas. Tal pressuposição é razoável para o estudo de fenômenos que envolvem a força de atitudes ou opiniões. O desenvolvimento econômico parece ser um desses. Todavia, para variáveis nominais, como raça ou gênero, por exemplo, a análise fatorial não faz sentido. Em última instância, a “propriedade” das técnicas disponíveis para estimar modelos de análise fatorial envolvendo variáveis categóricas depende da disposição do pesquisador para acreditar na existência de relação entre a variável observada, a variável medida e a variável latente, que se pretende medir (UCLA, 2004).

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Anderson, T. W. *An introduction to multivariate statistical analysis*. 2. ed. Toronto: John Wiley & Sons, 1984.
- Bartoli, H. *A economia serviço de vida: crise do capitalismo, uma política de civilização*. Lisboa: Instituto Piaget, 1996.
- Braga, T. M.; Freitas, A. P. G. de; Duarte, G. de S.; Carepa-Sousa, J. *Índices de sustentabilidade municipal: o desafio de mensurar*. Belo Horizonte: UFMG/Cedeplar, 2003. (Texto para discussão ; 225)
- Edwards Jr., C. H., Penney, D. E. *Calculus with analytic geometry*. 4th. ed. Englewood Cliffs, New Jersey: Prentice Hall International, 1994.
- Johnson, R. A., Wichern, D. W. *Applied multivariate statistical analysis*. 2. ed. Englewood Cliffs, New Jersey: Prentice Hall International, 1982.
- Kubrusly, L. S. Um procedimento para calcular índices a partir de uma base de dados multivariados. *Pesquisa Operacional*, **21**(1):107-117, jun. 2001.
- Mardia, K. V., Kent, J. T., Bibby, J. M. *Multivariate analysis*. London, New York: Academic Press, 1979.
- Parris, T. M.; Kates, R. W. Characterizing and measuring sustainable development. *Annual Review of Environmental Resources*, 28:559–86, 2003.
- Rangan, H. Political ecology and regional sustainability: reflections on contemporary debates and material practices. In: Gale, F. P.; M’Gonigle, R. M. (eds.) *Nature, production, power: towards an ecological political economy*. Cheltenham, U.K.; Northampton, USA: Edward Elgar, 2000.
- Siqueira, T. V. de; Siffert Filho, N. F. Desenvolvimento regional no Brasil: tendências e novas perspectivas. *Revista do BNDES*, **8**(16):79-117, dez. 2001.

SPSS for Windows. *Tutorial: version 11.0.0.* Chicago: Lead Technologies Inc., 2001.

Tragtenberg, M. Apresentação. In: Weber, M. *Weber.* São Paulo: Nova Cultural, 1997. (Coleção Os Economistas)

UCLA.University of California Los Angeles. *Lecture notes.*
<http://www.gseis.ucla.edu/courses/ed231a1/lect.html>, 2004.