

DESENVOLVIMENTO CONVERGENTE OU DIVERGENTE ENTRE OS ESTADOS BRASILEIROS? UMA ANÁLISE MULTIDIMENSIONAL, 1970-2000¹

André M. Marques²

Adelar Fochezatto^{3,4}

RESUMO

Explorando um banco de dados multidimensional para os estados brasileiros, em quatro momentos do tempo, o objetivo principal do estudo é verificar se a distância entre as unidades federativas vem aumentando ou diminuindo nos últimos quarenta anos em termos de desenvolvimento econômico. De forma complementar, através de análise de variância (MANOVA), o trabalho também testa as diferenças entre os grupos no período 1970-2000. A metodologia empregada baseou-se na Análise de Cluster para detectar o padrão de convergência ou divergência entre os estados e na estatística Λ de Wilks para testar a significância das diferenças entre os grupos. O teste de Wilcoxon foi empregado para testar a igualdade das distâncias entre os estados ao longo do tempo. A partir dos resultados da análise de Cluster foi possível constatar que o País está diminuindo a heterogeneidade *dentro* de dois grupos distintos de estados, porém, há um aumento das distâncias *entre* esses grupos de estados. Ao final, dois grupos de estados se configuram no Brasil: as macrorregiões Sul-Sudeste e Norte-Nordeste. Há evidência de que ambas estão convergindo internamente em várias dimensões do desenvolvimento, porém, enquanto grupo, ambas estão divergindo ao longo do tempo.

Palavras-chave: Convergência; divergência; desenvolvimento econômico.

ABSTRACT

Exploring a multidimensional database for the Brazilian states, in four time periods, the main objective of the study is to verify that the distance between the federative units is increasing or decreasing in the last forty years in terms of economic development. In a complementary way, using analysis of variance (MANOVA), the study also tests the differences between groups in the period 1970-2000. The methodology used was based on cluster analysis to detect the pattern of convergence or divergence between the states and the Wilks statistic to test the significance of

1

O primeiro autor agradece o apoio financeiro da Capes (AEX 11426/12-8) para a apresentação do trabalho no III Congresso Latinoamericano de Historia Económica, San Carlos de Bariloche (AR), de 23 a 27 de outubro de 2012.

² Doutor em Economia. Professor Adjunto da UFRN. Endereço eletrônico: 00094751@ufrgs.br

³ Doutor em Economia. Professor Titular da PUCRS. Pesquisador do CNPq. Endereço eletrônico: adelar@pucrs.br

⁴ Os autores agradecem os vários comentários e sugestões do Prof. Dr. Erik A. de Figueiredo em uma versão anterior do trabalho. Erros e omissões que persistam são de responsabilidade dos autores.

differences between groups. The Wilcoxon test was used to test for equality of the distances between the states over time. From the results of cluster analysis it was found that the country is decreasing heterogeneity in two distinct groups of states, however, there is an increase in distance between these groups of states. Finally, two groups of states are configured in Brazil: the macro-south-southeast and north-northeast. There is evidence that both are converging internally in various dimensions of development, however, as a group, both are diverging over time.

Key-words: Convergence; divergence; economic development.

JEL: O11; O18; O47.

1. Introdução

A hipótese da convergência (condicionada ou incondicionada) entre países ou regiões vem recebendo especial atenção dos pesquisadores desde meados dos anos 90. A importância dessa suposição reside precisamente na possibilidade teórica de que países ou regiões podem vir a alcançar um mesmo patamar de desenvolvimento em algum ponto do tempo, mesmo que iniciem seu processo de desenvolvimento em momentos diferentes.

A imensa maioria dos trabalhos que investiga essa suposição dá grande ênfase ao emprego de apenas um indicador de desenvolvimento, normalmente o PIB per capita. Isto porque, em princípio, essa seria uma medida de bem estar global da sociedade. Mais especificamente, essa seria uma medida da oferta disponível de bens e serviços por pessoa da comunidade. Teoricamente, Galor (1996) define três tipos de convergência de renda: a absoluta, segundo a qual a renda per capita dos países converge no longo prazo independentemente das condições iniciais; a condicionada, em que a renda per capita converge na medida em que os países tenham as mesmas características estruturais (preferências, tecnologia, taxa de crescimento da população, política econômica, etc.) independente das condições iniciais; e, finalmente, a hipótese dos clubes de convergência, mediante a qual ocorre polarização, pobreza e desigualdade persistentes e formação de clusters de crescimento. Galor (1996) mostra que as três hipóteses são competitivas entre si e podem ser obtidas como resultados analíticos do modelo de Solow tradicional a partir da inclusão de diferentes suposições teóricas acerca de capital humano, distribuição de renda, fertilidade, padrão de concorrência entre as empresas, etc.

Todavia, empiricamente, uma limitação básica dos trabalhos que investigam a hipótese de convergência é o emprego do PIB per capita como medida única, haja vista que ela não informa sobre a *utilização* dos bens e serviços da economia, apenas informa a sua disponibilidade. Por exemplo, pode ser que o PIB per capita seja bastante elevado em certo país ou estado, entretanto, neste mesmo local a renda e a propriedade sejam tão concentradas que a maior parte da população se aproprie de uma parcela demasiadamente pequena da oferta total de bens e serviços. Isto é, apesar da disponibilidade, uma parte da população pode não ter acesso aos bens e serviços produzidos. Isto pode se refletir nos indicadores educacionais (baixa escolaridade, analfabetismo, etc.), posto que o serviço de educação não seja acessado por todos igualmente, e, também, nos indicadores de saúde, habitação, saneamento, etc, afetando assim, o desenvolvimento econômico do país ou estado em consideração.

Por esses motivos, o objetivo do presente trabalho é verificar se está ou não havendo convergência de desenvolvimento entre os estados brasileiros usando uma abordagem multidimensional. A idéia básica é considerar outras dimensões do desenvolvimento, que informem, além da oferta de bens e serviços por pessoa na indústria e na agropecuária, também o uso dessa oferta de bens e serviços. Vários indicadores de desenvolvimento econômico são empregados na análise procurando capturar as múltiplas dimensões em que o desenvolvimento pode se manifestar, incluindo condições de moradia, desenvolvimento industrial, pobreza e desigualdade, renda, longevidade e educação, oferta de alimentos e nível de emprego.

A metodologia empregada para detectar o padrão de convergência ou divergência multidimensional entre os estados baseia-se na Análise de Agrupamentos (*Cluster Analysis*). Para testar a significância das diferenças entre os grupos formados, utiliza-se a estatística Λ de Wilks e, para testar a igualdade das distâncias entre os estados ao longo do tempo, emprega-se o teste de Wilcoxon.

O artigo possui cinco seções e está organizado do seguinte modo: após esta Introdução, a seção dois faz uma breve revisão da literatura; a seção três apresenta a metodologia e a base de dados; a seção quatro analisa os resultados encontrados; e, por fim, na seção cinco, são feitos os comentários finais.

2. Revisão da literatura

O estudo do desenvolvimento econômico em sua dimensão espacial tem sido dominada por duas perspectivas contrastantes. A primeira, com um enfoque mais tradicional, sustenta que o desenvolvimento tende a ser um processo uniforme no espaço, desde que não haja barreiras à operação das forças de mercado: em uma economia nacional, completamente integrada, espera-se uma convergência de renda e desenvolvimento ao longo do tempo pelo movimento dos preços, salários e taxa de lucro. Diferenças nas rendas colocariam em movimento mecanismos de correção das disparidades que levam à convergência entre regiões de um espaço nacional.⁵ Por isso, nessa perspectiva, disparidades regionais não seriam persistentes, pois a sua ocorrência seria compensada pela operação do mecanismo de mercado, corrigindo os desequilíbrios regionais. Entretanto, como se observou anteriormente, ainda que formulada sobre bases mais tradicionais, a hipótese dos clubes de convergência tem ganhado respaldo tanto na literatura teórica⁶ quanto empírica.⁷ Neste caso, o modelo tradicional de Solow poderia gerar a formação de clusters e a persistência de desigualdades de renda per capita e desenvolvimento mesmo no longo prazo.

De acordo com a perspectiva alternativa, sintetizada por Martin e Sunley (1998), não se pode esperar teoricamente que a convergência de renda e desenvolvimento entre regiões seja o caso mais provável, mesmo para longos períodos de tempo, dentro de um mesmo país. Para autores como Myrdal (1968), Kaldor (1978), Dixon e Thirlwall (1975), a divergência de renda e desenvolvimento entre as regiões é o caso mais provável. A explicação teórica baseia-se no princípio da causalção cumulativa, em que as forças de mercado, quando atuam livremente, tendem a reforçar as diferenças regionais. Economias de escala e de aglomeração levam a uma cumulativa concentração de investimentos em capital fixo, trabalho qualificado e alta produtividade em certas regiões pólo, em detrimento de outras regiões menos desenvolvidas. Mesmo com a ação política de governos no sentido da desconcentração da produção e dos investimentos, a convergência é uma hipótese improvável de ocorrer.⁸ Setterfield (1997) desenvolve um modelo teórico na perspectiva de Kaldor (1978), mostrando a possibilidade teórica de divergência persistente entre países ou regiões⁹.

Quando se considera o modelo de Solow como base teórica para a hipótese da convergência, questões metodológicas complicadas emergem, pois é necessário postular que os casos considerados já se encontram em seu estado estacionário (Ferreira e Cruz, 2008). A suposição teórica que fundamenta a hipótese de convergência é a de produtividade marginal decrescente do capital. De acordo com essa suposição, regiões com reduzido estoque de capital teriam altas taxas de retorno e, deste modo, desenvolver-se-iam mais rapidamente que as regiões ricas, que têm um estoque de capital mais elevado e retorno mais baixo. A falácia do argumento consiste em supor que uma alta taxa de retorno potencialmente considerada seja suficiente para efetivar uma alta taxa de

5

Ver Martin e Sunley (1998: 201) para referências a esse respeito. A possibilidade de divergência também pode ser deduzida dos modelos da “nova” teoria do crescimento, como o modelo AK, porém, por diferentes motivos teóricos daqueles apresentados neste estudo.

⁶ Ver, especialmente, Galor (1996).

⁷ Ver, especialmente, Durlauf e Johnson (1995) e Quah (1996) para suporte empírico a essa hipótese.

⁸ Por exemplo, a migração é sempre seletiva no fator idade; os jovens em geral migram para regiões mais desenvolvidas à custa de sua região de origem que perde capacidade de trabalho com a migração. Sobre esse aspecto, após analisar vários estudos na área, Martin e Sunley (1998: 212) concluem: “These analyses suggest that the migration of better-educated, highly skilled, and more enterprising workers is regionally disequilibrating, in that it benefits destination regions at the expense of the areas of origin”.

⁹ A hipótese de raiz unitária é consistente com os resultados teóricos do modelo kaldoriano uma vez que implica uma trajetória *path dependent* (ver Setterfield (1997)). Referindo-se a uma grande quantidade de estudos que testam a hipótese de raiz unitária Granger conclui: “The presence of a GUR [Generic Unit Root] in many series is a clear empirical fact which can not be ignored by economic theory, although the theory could be helpful in limiting the classes of models considered to explain this fact” (Granger, 1993, p. 310).

Sítio web: <http://www.aahce.fahce.unlp.edu.ar/Jornadas/iii-cladhe-xxiii-jhe/>

San Carlos de Bariloche, 23 al 27 de Octubre de 2012 - ISSN 1853-2543

acumulação de capital. De fato, uma alta taxa de retorno só poderá ser obtida *depois* que uma alta taxa de acumulação for concretizada pelos capitalistas e/ou pelo estado.¹⁰

A experiência tem mostrado que, em geral, as regiões mais desenvolvidas é que se tornam alvo dos investimentos de empresas e acumulação de capital fixo, em vista da mão de obra já qualificada, da infra-estrutura de transporte e comunicações, facilidades locais, etc. Ao final, a acumulação de capital fixo em certas áreas reforça os desequilíbrios regionais de uma forma acumulativa: quanto mais desenvolvida uma região, mais desenvolvida tende a ser, e vice-versa.

Uma das idéias chave do trabalho de Myrdal (1968), Kaldor (1978), Dixon e Thirlwall (1975) consiste na lei de Verdoon. Na medida em que uma região obtém uma vantagem comparativa para seu crescimento ela tende a mantê-la ou aumentá-la ao longo do tempo através dos retornos crescentes da indústria e seus efeitos de encadeamento sobre o produto agregado. Segundo Kaldor (1978: 141): “The prevailing distribution of real income in the world – the comparative riches or poverty of nations, or regions – is largely to be explained, not by ‘natural factors’, but by unequal incidence of development in industrial activities”.

Muitos trabalhos têm sido realizados para investigar a hipótese da convergência ou divergência no crescimento econômico. A hipótese da convergência tem recebido maior atenção dos pesquisadores. Ferreira e Cruz (2008), com uma base de dados municipal para os períodos de 1991 e 2000, utilizando um modelo de “regressão em árvore” com um parâmetro de *threshold*, investigaram a hipótese de convergência de renda dos municípios brasileiros. O objetivo do trabalho era verificar se houve convergência na desigualdade entre os municípios e se as condições iniciais, medidas pelo índice de Gini, geram diferentes equilíbrios (clubes de convergência). Os autores constataam a ocorrência de seis clubes de convergência, em que a renda do trabalho é de longe a maior influência para a redução das desigualdades. As transferências governamentais e a escolaridade contribuem pouco para esse processo.

Gligor e Ausloos (2011) investigaram a hipótese de convergência sob um enfoque multidimensional a partir de dados de séries temporais, considerando variáveis como PIB, PIB per capita, consumo e investimento para o continente europeu. Através de análise de cluster e correlação os autores concluem que há evidência de formação de clubes de convergência, dentro de uma mesma região da Europa. Para os autores, as aglomerações industriais e o crescimento são forças que se retroalimentam no sentido da formação dos clubes de convergência.

Marques (2009), empregando uma base de dados municipal para os anos 1990 e 1999, através de análise de variância e correlação, definindo *dummies* regionais, constatou que as três macrorregiões que constituem o estado do Rio Grande do Sul diferem significativamente em termos de crescimento e desigualdade, com uma significativa correlação inversa entre ambas. Segundo o autor, os municípios situados na região nordeste do estado combinam mais alto crescimento com um grau de desigualdade substancialmente menor. Embora o período de uma década seja relativamente curto, o trabalho conclui que os resultados apontam para uma divergência entre as regiões do Rio Grande do Sul.

Moreira et. al. (2010), empregando uma base de dados estadual para o período de 1996 a 2007, através de modelos dinâmicos de equações simultâneas para dados em painel, constataram a persistência de disparidades regionais no Brasil e concluem que essa segmentação tem efeito desacelerador sobre as taxas de investimento produtivo a longo prazo, portanto, sobre as taxas de crescimento econômico, perpetuando a incidência da pobreza através de um círculo vicioso ao longo do tempo, de modo que a coexistência da pobreza e desigualdade dificulta o crescimento econômico. Segundo os autores, melhorar a equidade em um ambiente marcado pela alta heterogeneidade teria o efeito de acelerar a taxa de crescimento a longo prazo. Este, por sua vez, atuaria sobre a redução da pobreza, levando a um maior nível de equidade e maior crescimento,

¹⁰ Teoricamente, os capitalistas ganham o que gastam.

engendrando um círculo virtuoso de desenvolvimento. Os autores ressaltam a ocorrência de grandes disparidades entre os estados situados nas regiões Norte/Nordeste vis-à-vis região Sul.

Na mesma perspectiva teórica de Myrdal (1968), Aryeetey et. al. (2009), utilizando o Gini da renda e três coeficientes de entropia generalizada, analisam a existência de disparidades regionais em Gana. A base de dados leva em conta o Gini da renda e vários indicadores de desenvolvimento, como o consumo de energia elétrica, acesso à escola, taxa de alfabetização, consumo de água potável (entre outros), para os períodos de 1991-2006 e 1970-2000. Os autores concluem que a persistente divergência entre as regiões do país (norte-sul; rural-urbano) tem suas origens no processo de colonização, porém, as condições foram reforçadas pelas estratégias de política que se seguiram. Por exemplo, o processo de substituição de importações e expansão dos investimentos nas áreas de exportação (indústria de mineração) criou infra-estrutura e facilidades de transporte em áreas concentradas, reforçando as disparidades entre as regiões, em especial nas zonas portuárias, em detrimento de áreas menos desenvolvidas. Nessas áreas, criou-se um incentivo para novos investimentos e deslocamento de mão de obra qualificada, em detrimento das demais regiões.

O presente trabalho, através de análise de cluster, tem por objetivo verificar, em quatro momentos do tempo, se existe algum padrão de convergência ou divergência em termos de desenvolvimento econômico. Especificamente, o objetivo é procurar capturar, de forma sintética, uma medida de separação ou aproximação ao longo dos anos que leve em conta uma grande quantidade de variáveis que expressam desenvolvimento econômico (quatorze variáveis), e não apenas o PIB per capita, tradicionalmente empregado na literatura. A escolha da metodologia é adequada a essa temática, pois permite a sintetização de uma grande quantidade de variáveis medidas nos casos (estados) em uma única matriz de distâncias, que, por fim, resulta em uma árvore de agrupamentos hierárquicos, dos mais semelhantes para os mais dessemelhantes. Uma abordagem similar foi empregada por Saint-Arnaud e Bernard (2003), para verificar se estaria ocorrendo convergência entre os regimes de bem-estar de diferentes países.

3. Metodologia e base de dados

3.1 Descrição dos dados e software

Os dados usados no trabalho estão no Quadro 1. Para facilitar a análise, essas variáveis podem ser agrupadas em seis dimensões do desenvolvimento: condições de moradia (krrpc, krupc, cerpc); desenvolvimento industrial (ceipc, pindpc); pobreza e desigualdade (theil, p0); desenvolvimento humano (idhl, idhe, idhr); produção agrícola (yappc, yatpc); e emprego (empurb, emprur).

Quadro 1: Descrição das variáveis utilizadas no trabalho

Variável	Descrição
krrpc	Capital residencial rural - per capita - Capital Residencial - Rural - R\$ de 2000 (mil)
krupc	Capital residencial urbano - per capita - Capital Residencial - Urbano - R\$ de 2000 (mil)
cerpc	Consumo de energia elétrica - residencial per capita - MWh
ceipc	Consumo de energia elétrica - industrial per capita - MWh
theil	Renda - desigualdade - índice L de Theil (1970)
p0	Pobreza - pessoas pobres (P0) - (%)
idhl	Índice de Desenvolvimento Humano (IDH) - longevidade
idhe	Índice de Desenvolvimento Humano (IDH) - educação
idhr	Índice de Desenvolvimento Humano (IDH) - renda
yappc	Produção - lavoura permanente - R\$ de 2000 (mil) - per capita
yatpc	Produção - lavoura temporária - R\$ de 2000 (mil) - per capita
empurb	População urbana ocupada como proporção da população total.

emprur	População rural ocupada como proporção da população total.
pindpc	PIB industrial estadual per capita - R\$ de 2000 - valor adicionado - preços básicos - R\$ de 2000 (mil)

Fonte: elaboração própria.

Notas: a) todas as variáveis foram obtidas no IPEADATA (<http://www.ipeadata.gov.br>); b) não foram encontrados dados para consumo de energia elétrica em Mato Grosso do Sul no ano de 1971 e em Tocantins nos anos de 1971 e 1980; e c) não foram encontrados dados para o PIB industrial para Mato Grosso do Sul e para Tocantins em 1971. Em vista disso, essas duas unidades da federação não foram incluídas no estudo.

Esses dados foram organizados para o formato de matriz ($n \times p$), em que n é o número de indivíduos (casos) e p é o número de variáveis medidas. Neste caso, os casos são as unidades da federação e as variáveis as diferentes dimensões do desenvolvimento econômico, tais como estoque de capital físico per capita, desigualdade de renda, PIB industrial per capita e assim sucessivamente. A descrição completa das variáveis encontra-se no Apêndice. Todas as estimações e gráficos foram feitos no Ambiente **R**.

No total, quatro matrizes de dados foram organizadas para os últimos quarenta anos da economia brasileira, em uma base estadual: 1970; 1980; 1991 e 2000. Portanto, as dimensões do desenvolvimento são avaliadas em quatro momentos do tempo para todos os estados brasileiros, excetuando-se Tocantins e Mato Grosso do Sul.

A matriz de correlação amostral com os coeficientes de Pearson ($R_{14 \times 14}$) oferece uma indicação do relacionamento linear entre as seis dimensões do desenvolvimento econômico. O objetivo da análise de correlação, preliminarmente, é detectar padrões de relacionamento entre as variáveis analisadas, que, em princípio mudam pouco com o tempo. O Quadro 2 apresenta o coeficiente de correlação amostral para cada par de variáveis do estudo nos anos de 1970, 1980, 1991 e 2000. Espera-se, teoricamente, que certos padrões de relacionamento se mantenham estáveis ao longo do tempo, tais como a relação negativa entre a percentagem de pobres e os IDH's da renda, longevidade e educação.

A análise das matrizes de correlação permite que sejam extraídas algumas conclusões preliminares sobre o relacionamento entre as variáveis analisadas. Primeiro, a melhora nas condições de moradia, ao longo das quatro décadas, tanto urbanas quanto rurais, está altamente correlacionada com melhora nos IDH's (educação, renda e longevidade), cujo coeficiente de correlação é, em geral, maior que 0,70 e em alguns casos maior que 0,80. Ademais, as melhores condições de moradia estão fortemente relacionadas à redução da percentagem de pobres. Em relação à melhora nas condições de moradia, em especial, verifica-se que são as melhores condições de moradia do meio rural (e não as urbanas) que mais fortemente se relacionam com a redução do grau de desigualdade da renda.

Segundo, a produção agrícola per capita, ao longo das quatro décadas foi deixando de apresentar uma correlação importante com os indicadores de desenvolvimento humano (IDH's). Isto pode ser observado nas duas últimas colunas das matrizes de correlação. No ano de 1970, por exemplo, a produção agrícola per capita apresentava uma correlação muito alta com a redução da percentagem de pobres (-0,94) e também com a melhora dos IDH's. Entretanto, nos anos 1980, 1991 e 2000 essa correlação praticamente deixou de existir.

Terceiro, em contraste com a produção agrícola per capita, a produção industrial per capita manteve ao longo dos anos sua forte influência no sentido da redução da percentagem de pobres e aumento dos IDH's. Além disso, a produção industrial per capita, ao longo dos anos, passou a ter importante influência no sentido da redução da desigualdade, medida pelo índice de Theil, algo que também não se observa no caso da produção agrícola. Esta redução da desigualdade pode estar relacionada com os ganhos de produtividade ocasionados pela participação da indústria e as correspondentes melhorias salariais. Os resultados obtidos por Ferreira e Cruz (2008), com relação à renda do trabalho e sua influência sobre a redução da desigualdade corroboram essa interpretação.

Quarto, com relação à utilização da mão de obra nos setores urbano e rural da economia, constata-se um vivo contraste entre esses dois setores e sua marcada influência sobre as condições de moradia, os IDH's, a pobreza e a desigualdade, especialmente em 1991 e 2000. Em particular, a

taxa de ocupação da mão de obra no setor urbano tem forte correlação positiva com a melhora nas condições de moradia, elevação dos IDH's; por outro lado, a utilização da mão de obra no setor rural provoca fortes efeitos contrários.

	<i>krnpc</i>	<i>krupc</i>	<i>cerpc</i>	<i>ceipc</i>	<i>theil</i>	<i>p0</i>	<i>idhl</i>	<i>idhe</i>	<i>idhr</i>	<i>yappc</i>	<i>yatpc</i>	<i>empurb</i>	<i>emprur</i>	<i>pindpc</i>
krnpc	1,00													
krupc	0,73	1,00												
cerpc	0,73	0,85	1,00											
ceipc	0,46	0,67	0,67	1,00										
theil	-0,27	0,18	-0,03	0,23	1,00									
p0	-0,88	-0,86	-0,89	-0,54	0,19	1,00								
idhl	0,71	0,69	0,53	0,38	-0,12	-0,57	1,00							
idhe	0,87	0,85	0,76	0,50	-0,08	-0,84	0,83	1,00						
idhr	0,87	0,88	0,92	0,57	-0,08	-0,98	0,59	0,84	1,00					
yappc	-0,51	-0,27	-0,29	-0,11	0,47	0,38	-0,50	-0,41	-0,36	1,00				
yatpc	0,81	0,78	0,85	0,52	-0,11	-0,94	0,42	0,67	0,96	-0,36	1,00			
empurb	0,73	0,88	0,93	0,60	0,17	-0,89	0,47	0,76	0,94	-0,19	0,88	1,00		
emprur	-0,60	-0,73	-0,86	-0,47	-0,18	0,79	-0,31	-0,67	-0,82	0,07	-0,75	-0,93	1,00	
pindpc	0,65	0,59	0,77	0,82	-0,10	-0,65	0,43	0,57	0,65	-0,18	0,60	0,62	-0,54	1,00

1970

	<i>krnpc</i>	<i>krupc</i>	<i>cerpc</i>	<i>ceipc</i>	<i>theil</i>	<i>p0</i>	<i>idhl</i>	<i>idhe</i>	<i>idhr</i>	<i>yappc</i>	<i>yatpc</i>	<i>empurb</i>	<i>emprur</i>	<i>pindpc</i>
krnpc	1,00													
krupc	0,83	1,00												
cerpc	0,68	0,84	1,00											
ceipc	0,68	0,62	0,33	1,00										
theil	-0,38	0,06	-0,08	-0,17	1,00									
p0	-0,86	-0,83	-0,82	-0,49	0,31	1,00								
idhl	0,65	0,51	0,46	0,35	-0,47	-0,78	1,00							
idhe	0,84	0,79	0,76	0,45	-0,27	-0,96	0,81	1,00						
idhr	0,74	0,72	0,68	0,46	-0,21	-0,95	0,79	0,92	1,00					
yappc	0,07	0,08	-0,05	0,47	0,16	-0,04	0,11	0,01	0,12	1,00				
yatpc	0,21	0,01	-0,15	0,01	-0,15	-0,24	0,41	0,22	0,38	-0,03	1,00			
empurb	0,83	0,96	0,88	0,54	0,03	-0,88	0,50	0,85	0,79	0,00	0,06	1,00		
emprur	-0,67	-0,84	-0,82	-0,46	-0,09	0,74	-0,26	-0,72	-0,63	0,09	0,23	-0,90	1,00	
pindpc	0,84	0,67	0,53	0,70	-0,41	-0,75	0,56	0,70	0,65	0,16	0,15	0,67	-0,52	1,00

1980

	<i>krnpc</i>	<i>krupc</i>	<i>cerpc</i>	<i>ceipc</i>	<i>theil</i>	<i>p0</i>	<i>idhl</i>	<i>idhe</i>	<i>idhr</i>	<i>yappc</i>	<i>yatpc</i>	<i>empurb</i>	<i>emprur</i>	<i>pindpc</i>
krnpc	1,00													
krupc	0,86	1,00												
cerpc	0,84	0,88	1,00											
ceipc	0,48	0,29	0,17	1,00										
theil	-0,63	-0,28	-0,51	-0,43	1,00									
p0	-0,79	-0,75	-0,93	-0,12	0,63	1,00								
idhl	0,76	0,73	0,83	0,16	-0,60	-0,91	1,00							
idhe	0,77	0,74	0,89	0,16	-0,58	-0,96	0,92	1,00						
idhr	0,78	0,80	0,94	0,06	-0,48	-0,98	0,87	0,94	1,00					
yappc	0,02	-0,06	0,00	0,01	-0,06	-0,05	0,04	0,13	0,03	1,00				
yatpc	-0,01	-0,12	-0,02	-0,19	-0,27	-0,16	0,13	0,12	0,08	0,37	1,00			
empurb	0,80	0,90	0,91	0,07	-0,44	-0,88	0,82	0,84	0,89	-0,03	0,04	1,00		
emprur	-0,58	-0,67	-0,75	0,05	0,23	0,67	-0,56	-0,61	-0,72	0,21	0,35	-0,80	1,00	
pindpc	0,58	0,64	0,67	0,40	-0,47	-0,60	0,61	0,58	0,56	-0,01	-0,08	0,58	-0,44	1,00

1991

(continua)

	<i>krpc</i>	<i>krupc</i>	<i>cerpc</i>	<i>ceipc</i>	<i>theil</i>	<i>p0</i>	<i>idhl</i>	<i>idhe</i>	<i>idhr</i>	<i>yappc</i>	<i>yatpc</i>	<i>empurb</i>	<i>emprur</i>	<i>pindpc</i>
krpc	1,00													
krupc	0,95	1,00												
cerpc	0,85	0,82	1,00											
ceipc	0,29	0,33	0,15	1,00										
theil	-0,50	-0,38	-0,62	-0,33	1,00									
p0	-0,85	-0,80	-0,92	-0,23	0,79	1,00								
idhl	0,72	0,69	0,77	0,29	-0,67	-0,87	1,00							
idhe	0,77	0,72	0,90	0,18	-0,71	-0,93	0,84	1,00						
idhr	0,91	0,88	0,96	0,18	-0,64	-0,97	0,84	0,92	1,00					
yappc	0,08	0,13	0,04	0,37	-0,36	-0,25	0,15	0,17	0,17	1,00				
yatpc	0,10	0,01	0,16	-0,02	-0,29	-0,32	0,31	0,23	0,24	-0,03	1,00			
empurb	0,87	0,85	0,93	0,20	-0,66	-0,96	0,83	0,90	0,96	0,13	0,29	1,00		
emprur	-0,56	-0,58	-0,79	0,07	0,22	0,60	-0,53	-0,67	-0,69	0,27	-0,02	-0,72	1,00	
pindpc	0,54	0,56	0,62	0,50	-0,51	-0,61	0,63	0,56	0,58	0,16	0,04	0,57	-0,37	1,00

Figura 1: Matrizes de Correlação dos anos de 1970, 1980, 1991 e 2000.

Fonte: elaboração própria.

Conclui-se que o papel de destaque para a indústria e sua contribuição para o desenvolvimento econômico, enfatizado por Kaldor (1978), nas quatro décadas analisadas, foi preponderante no caso do Brasil em termos de redução da percentagem de pobres, da desigualdade de renda medida pelo índice de Theil e aumento nos IDH's da renda, longevidade e educação. Por ser tipicamente um setor em que a intensidade da demanda se manifesta nos crescentes ganhos de produtividade (e na qualificação dos trabalhadores) e escala da produção, esses resultados sugerem que a indústria continua tendo papel preponderante no desenvolvimento econômico do Brasil. Algo que não se constata comumente na agricultura, considerada uma atividade sujeita a retornos decrescentes à escala (Thirlwall, 2005).

3.2 Análise de Cluster

O passo seguinte do trabalho consiste em verificar se os estados brasileiros situados em um grupo (cluster) de baixo desenvolvimento tiveram sua posição alterada ao longo dos últimos quarenta anos: ou se aproximaram dos mais desenvolvidos ou se afastaram deles, configurando algum padrão de convergência ou divergência. O primeiro caso configura uma situação de convergência, respaldando a hipótese daqueles que acreditam que o processo de desenvolvimento naturalmente leva a uma situação mais igualitária. Por outro lado, o segundo caso respalda a hipótese daqueles que vêem esse processo como sendo de natureza divergente. Para testar essas hipóteses, o primeiro passo consiste na separação inicial dos grupos.

A evidência desse comportamento será a mudança de um cluster de baixo desenvolvimento para um cluster de mais alto grau de desenvolvimento e o teste de Wilcoxon para a diferença das distâncias ao longo dos anos. Há dois casos extremos que servem de guia para a análise: o caso em que todos os estados têm o mesmo grau de desenvolvimento e estão todos situados no mesmo cluster e, o caso em que há grande dispersão em termos de desenvolvimento e há a formação de certo número de clusters hierárquicos. É importante verificar se houve mudança na composição desses clusters ao longo do período analisado ou se os estados inicialmente menos desenvolvidos continuam, ao longo do tempo, na mesma situação relativa.

Na identificação dos clusters o passo inicial consiste na definição da medida de distância a ser utilizada. Neste aspecto, a maioria dos autores emprega a medida de distância mais simples, a distância euclidiana. Para tanto, considere o vetor aleatório de medidas X_j definido por:

(1)

$$X_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

onde X_{ij} representa o valor observado da variável i medida no elemento j . Como a medida de distância reflete dissimilaridade, quanto menor seu valor, mais similares serão os casos que estão sendo comparados. A distância euclidiana entre dois elementos amostrais X_l e X_k , $l \neq k$, é definida por:

(2)

$$D_{25 \times 25} = \begin{bmatrix} 0 & d_{12} & \dots & d_{125} \\ d_{21} & 0 & \dots & d_{225} \\ \vdots & \vdots & \ddots & \vdots \\ d_{251} & d_{252} & \dots & 0 \end{bmatrix}$$

Os resultados do cálculo das distâncias foram organizados em uma matriz de distâncias, $D_{25 \times 25}$, uma vez que a análise é feita para vinte e cinco estados da federação brasileira. A partir da obtenção de $D_{25 \times 25}$, é possível aplicar um critério de ligação entre os indivíduos para a formação do dendograma, que é um gráfico em que no eixo das abscissas aparecem os clusters e no eixo das ordenadas aparecem as distâncias (adimensionais) calculadas entre os elementos amostrais.

Por levar em consideração os vetores de médias amostrais, baseando-se nos princípios da análise de variância em cada passo do processo de aglomeração, o método de Ward tem recebido especial interesse. Mingoti (2005, Cap. 6) oferece uma discussão detalhada e uma comparação deste com os outros critérios de aglomeração. Neste trabalho, segue-se a abordagem de Saint-Arnaud e Bernard (2003), que também empregaram o critério de Ward para a obtenção do dendograma.

Para o cálculo das matrizes de distância todas as variáveis foram normalizadas, haja vista que as unidades de medida poderiam influenciar os resultados (Mingoti, 2005). A idéia subjacente à separação em grupos é de que a partição “desejada” ou esperada, resultante do cálculo das distâncias e do algoritmo de aglomeração, seja a que possua a maior heterogeneidade possível *entre* os grupos (clusters) formados e a maior homogeneidade possível *dentro* dos grupos.

A distância euclidiana e o método de Ward de aglomeração foram a base para a apresentação dos dendogramas para todos os anos considerados no estudo (Saint-Arnaud e Bernard, 2003; Carvalho et. al., 2006). O método de Ward procura, em cada passo da aglomeração, levar em conta estrutura de variabilidade dos dados, minimizando a perda de informação em cada passo do algoritmo (Johnson e Wichern, 1998: 751). A linha colorida que separa os clusters foi estabelecida de acordo com a sugestão de Pereira (2001), em que o maior salto no algoritmo de aglomeração determina a separação dos grupos. Os resultados da análise de cluster são apresentados e analisados na seção 4. Para a validação dos clusters identificados utilizou-se a estatística Λ de Wilks.

3.3 Validação dos resultados

A aplicação de técnicas multivariadas requer que as p variáveis em estudo sejam correlacionadas em um grau mínimo (Pereira, 2001, Mingoti, 2005). Em particular, a estatística Λ de Wilks depende da existência de correlações não nulas entre as variáveis analisadas (Reis, 2001: 169). Admitindo-se a suposição de que o vetor aleatório \mathbf{X} segue distribuição normal p -variada, pode-se testar a hipótese de que a matriz de correlação populacional é uma identidade, contra a hipótese alternativa bicaudal. Na prática, a não rejeição da hipótese nula indica que as variáveis são independentes, portanto, inadequadas para a análise de correlação. A estatística de teste é dada por:

(3)

$$T = \frac{1}{n} \frac{\lambda_1 \lambda_2 \dots \lambda_p}{\lambda_1 + \lambda_2 + \dots + \lambda_p}$$

onde T é a estatística de teste, n é o tamanho da amostra, p é o número de variáveis, $\hat{\lambda}_j$ são os autovalores estimados a partir da matriz de correlação amostral do vetor \mathbf{X} . Para n suficientemente grande, a estatística T tem uma distribuição qui-quadrado com $\frac{1}{2}p(p-1)$ graus de liberdade.

Ademais, a separação em diferentes grupos de estados com diferentes indicadores de desenvolvimento pode ser testada quanto à sua significância estatística. Em especial, o pesquisador pode estar interessado em testar a hipótese nula de que não há diferença entre os grupos

identificados. A estatística Λ de Wilks, a partir do quociente de verossimilhanças, não se encontra tabelada, mas pode ser aproximada por:

$$(4) \quad \frac{(pn-k-p+1)\Lambda}{p} \sim F$$

que segue aproximadamente uma distribuição F com $(pn-k-p+1)$ graus de liberdade, em que p é número de variáveis, k é o número de grupos e n o número de indivíduos. A estatística Λ é obtida de acordo com a expressão,

$$(5) \quad \Lambda = \frac{S}{\sum_{j=1}^k (1+\lambda_j)^{-1}}$$

em que λ_j são os autovalores do produto de matrizes $(W^{-1} \cdot B)$ e s é o número de autovalores da matriz resultante. A matriz W expressa os produtos cruzados entre os grupos e dentro dos grupos, definida por:

$$(6) \quad W = \sum_{j=1}^k \sum_{u=1}^{n_j} (X_{ju} - \bar{X}_j)(X_{ju} - \bar{X}_j)'$$

Por sua vez, a matriz B é definida como,

$$(7) \quad B = \sum_{j=1}^k n_j (\bar{X}_j - \bar{X})(\bar{X}_j - \bar{X})'$$

em que $\bar{X}_j = \frac{\sum_{u=1}^{n_j} X_{ju}}{n_j}$ é o vetor de médias amostrais para o grupo j , $\bar{X} = \frac{\sum_{j=1}^k \sum_{u=1}^{n_j} X_{ju}}{n}$ é o vetor de

médias amostrais para a totalidade da amostra e $n = \sum_{j=1}^k n_j$ é o tamanho total da amostra. Neste

contexto, as hipóteses nula e alternativa a serem testadas são as seguintes: $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ e $H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$.

Visando validar a detecção dos clusters pelos critérios escolhidos, procedeu-se ao teste para diferença entre os grupos, descrito anteriormente na seção 3.3. Os resultados são apresentados na tabela 1. Em todos os casos constata-se que os grupos diferem estatisticamente a 0,01 de significância.

Tabela 1: Resultados do teste para a diferença entre os grupos.

Ano	$\hat{\Lambda}$	Estatística F com $(pn-k-p+1)$ graus de liberdade	Probabilidade exata do teste
1970	0,0008074	21,98***	0,0000
1980	0,0001152	11,82***	0,0000
1991	0,0000500	16,09***	0,0000
2000	0,0179000	39,19***	0,0000

Fonte: elaboração própria.

Nota: (***) estatisticamente significativo a 0,01 de significância.

A tabela 2 abaixo apresenta os resultados do teste de Bartlett para a matriz de correlação populacional. Em todos os casos a hipótese nula de ausência de correlação pode ser rejeitada a 0,01 de significância.

Tabela 2: Resultado do Teste de Bartlett para a matriz de correlação populacional.

Ano	Estatística \hat{T}	Probabilidade exata do teste
1970	573,05***	0,0000
1980	468,81***	0,0000
1991	488,84***	0,0000
2000	515,07***	0,0000

Fonte: elaboração própria.

Nota: (***) estatisticamente significativo a 0,01 de significância.

Sobre a hipótese de convergência ou divergência, aludidas anteriormente, alguma evidência pode ser obtida pelo teste de Wilcoxon para a igualdade das distâncias, contra hipótese alternativa específica de divergência ou convergência. Isto é, pode-se verificar, com base nos dados deste estudo se, dinamicamente, a distância entre os estados vem aumentando, configurando certa divergência, ou vem diminuindo, configurando algum tipo de convergência (ou está estabilizada) em termos de desenvolvimento econômico, levando-se em consideração as seis dimensões do desenvolvimento vistas anteriormente.

Seguindo o trabalho de Gligor e Ausloos (2011: p. 12-13), a partir do cálculo das quatro matrizes de distâncias, as distâncias euclidianas foram organizadas em vetores, e, então, foi utilizado o teste de Wilcoxon (para dados pareados) para testar a hipótese de igualdade das distâncias ao longo dos anos. Os resultados dos testes de hipóteses encontram-se nas Tabelas 3 e 4. Os resultados apresentados consideram a hipótese nula como igualdade das distâncias; na hipótese alternativa têm-se os resultados para o caso unicaudal à esquerda (divergência), como postulado teoricamente por Myrdal (1968) e Kaldor (1978).

4. Resultados e discussão

Considerando as múltiplas dimensões do desenvolvimento contempladas neste trabalho, pode-se dizer que, no período de 1970 a 2000, o Brasil ficou mais desigual, comparando o grupo de estados do Norte e Nordeste com o grupo de estados do Sudeste, Sul e Centro-Oeste; e mais igual entre os estados dentro desses dois grupos. Considerando um ponto de corte em uma distância igual a 10 (eixo das ordenadas)¹¹, em 1970 o Brasil tinha três grupos de estados com desenvolvimento relativamente similar. O primeiro grupo (o mais homogêneo) era formado pelos estados de São Paulo, Rio de Janeiro e Distrito Federal. O segundo grupo era formado pelos demais estados do Sudeste (Minas Gerais e Espírito Santo), mais os estados do Sul, do Centro-Oeste e do Norte. O terceiro grupo era formado pelos estados do Nordeste.

Ao longo do tempo, houve uma divisão no segundo grupo: os estados do Norte convergiram para o terceiro grupo (Nordeste), enquanto que os demais estados do grupo convergiram para o primeiro grupo. Em momento algum, ao longo de todo o período, houve convergência entre os estados do primeiro e terceiro grupos. Pelo contrário, a distância entre eles aumentou no período. Os Anexos 1 e 2 apresentam, respectivamente, as distâncias multidimensionais entre todas as unidades da federação em cada período analisado e os valores médios das 14 variáveis consideradas no trabalho por cluster e por período. Essas informações são úteis para acompanhar a evolução das distâncias (diferenças) entre todas as unidades estaduais analisadas, verificar as diferenças entre os clusters formados e acompanhar a evolução dos níveis de desenvolvimento desses clusters.

Analisando mais detalhadamente os grupos formados ao longo do tempo, a Figura 1 mostra a posição dos estados brasileiros com relação às variáveis medidas em 1970. Consta-se que as distâncias entre os clusters não eram substanciais, entretanto, eram suficientes para gerar pelo menos três grupos de estados, em que os mais dessemelhantes em relação aos demais estão situados

¹¹ Os dendogramas das figuras 1 a 4 mostram, nos eixos das ordenadas, a distância euclidiana multidimensional entre os grupos de unidades da federação (distância considerando as 14 variáveis usadas como indicadores de desenvolvimento). Nos eixos das abscissas, os dendogramas mostram as Unidades da Federação e os grupos a que elas pertencem. A separação dos grupos foi validada pela estatística Λ de Wilks, cujos resultados aparecem na Tabela 1.

no cluster mais próximo da origem. Observa-se que a distância que separa esse cluster dos demais está próxima de 30. Enquanto os demais clusters estão separados por uma distância de dimensão 15.

Não obstante, observa-se que há grande homogeneidade dentro dos clusters, pois as distâncias que os separam, medidas no eixo das ordenadas, são muito pequenas comparadas com as distâncias que separam as aglomerações (saltos). Esse é um resultado desejável, do ponto de vista dos princípios da aglomeração.

Analisando-se a figura 2, constata-se que, em 1970, o Nordeste era formado por dois grupos distintos de estados, mais o Acre. Entretanto, a maior parte dos estados mesclava-se em um grande grupo que incluía estados do Norte (AM, RO) juntamente com estados do Centro-Oeste (MT, GO) e Sul (RS, SC, PR). Praticamente 50% dos estados brasileiros pertenciam a um único cluster de desenvolvimento, incluindo os estados do Sul, Centro-Oeste e Norte. DF, RJ e SP formam um grupo separado que se distingue fortemente dos demais, haja vista o tamanho da distância que os separa dos demais.

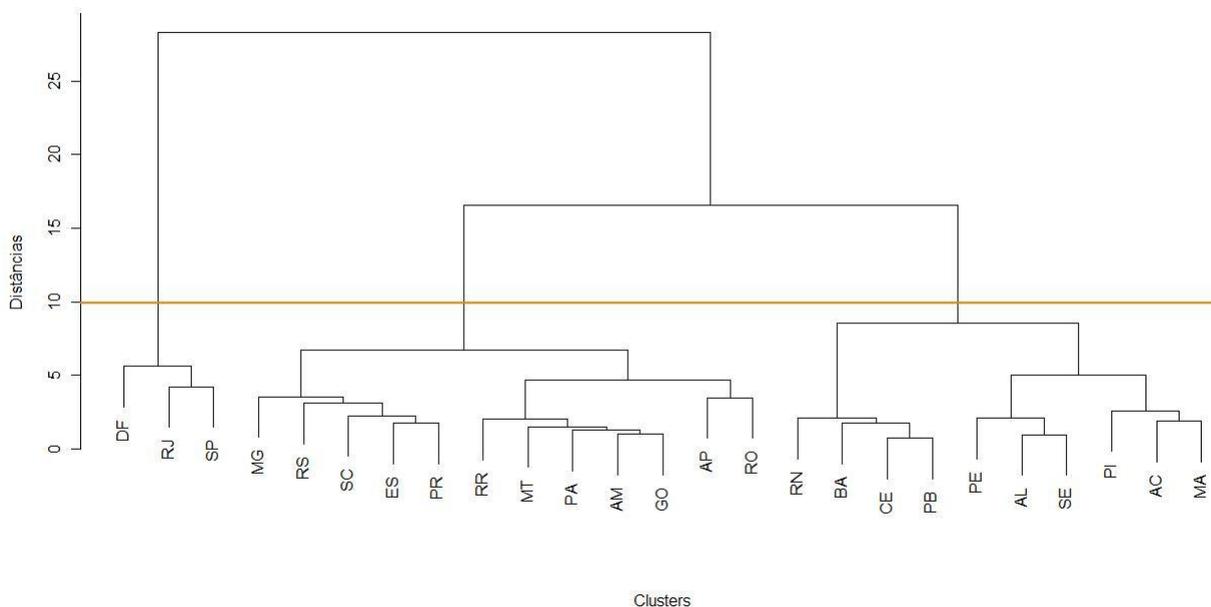


Figura 2: Dendrograma mostrando os resultados da aglomeração de unidades da federação, 1970.
Fonte: elaboração própria.

Na figura 3, tem-se a nova posição ocupada pelos estados brasileiros, uma década depois. Observa-se que houve uma considerável alteração na disposição dos clusters. Uma constatação importante é a persistência da diferenciação de São Paulo em relação aos demais estados, que forma um cluster quase sozinho, levando consigo RJ e DF, como em 1970.

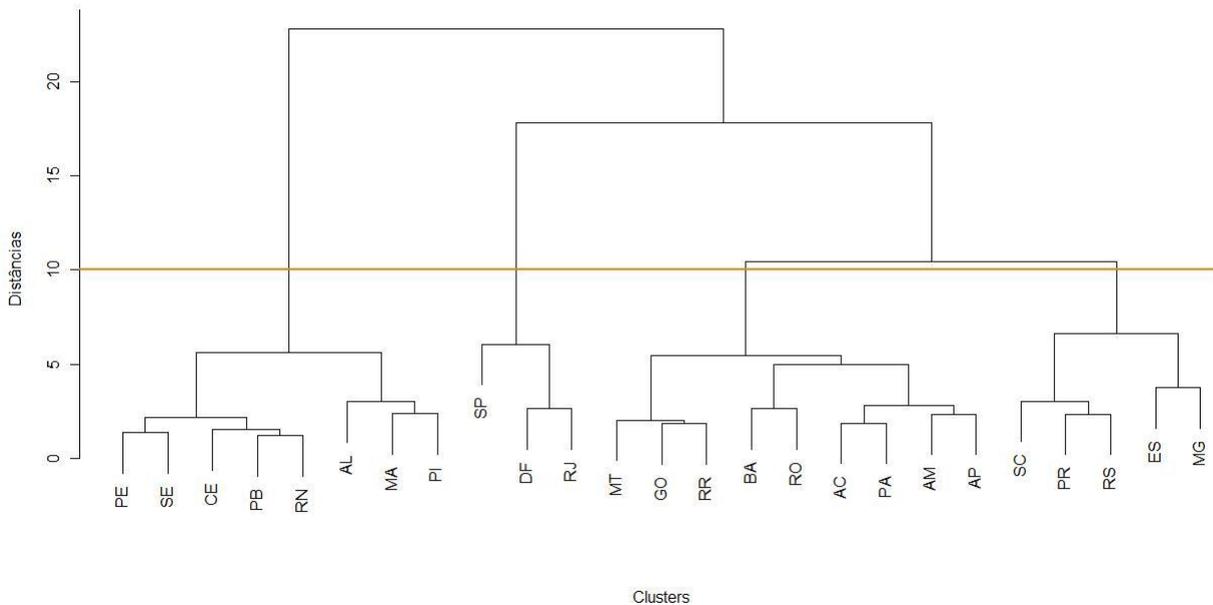


Figura 3: Dendrograma mostrando os resultados da aglomeração de unidades da federação, 1980.
 Fonte: elaboração própria.

Além da persistência do cluster formado por SP, RJ e DF, a região Nordeste (exceto BA) aparece visivelmente separada do Norte e Sul do Brasil. Em particular, os estados do Norte (como RO e RR) passam a formar um cluster com o Centro-Oeste (GO, MT). A região Sul, por sua vez, passa a formar um cluster com MG e ES. Constata-se que houve, portanto, grande alteração, em termos de desenvolvimento, entre os estados brasileiros, da década de 1970 para 1980.

A seguir, na figura 4, tem-se a nova posição ocupada pelos estados brasileiros, duas décadas depois, em 1991. Como se observou anteriormente, SP, RJ e DF continuam formando um cluster de desenvolvimento separado dos demais.



Figura 4: Dendrograma mostrando os resultados da aglomeração de unidades da federação, 1991.

Fonte: elaboração própria.

Em 1991, além de ES e MG, GO e MT passaram a fazer parte do cluster da região Sul, o que pode estar indicando sensível melhora nos seus índices de desenvolvimento, tendo em vista as médias das variáveis analisadas.

Outro aspecto notável é a nítida separação que passa a vigorar entre as regiões Norte e Nordeste do Brasil. Tem-se, portanto, a formação clara de quatro clusters distintos, com alto grau de similaridade entre os estados que os compõem, mas alta heterogeneidade entre os clusters, que sintetizam razoavelmente bem as regiões Norte, Nordeste, Sudeste e Sul do Brasil, em que a região Centro-Oeste aparece diluída na região Sul.

A figura 5 apresenta a configuração dos clusters para o ano 2000. Em relação ao ano de 1970, constata-se que houve a consolidação da separação dos estados brasileiros em termos de desenvolvimento econômico. Em particular, no ano 2000 os estados de RO, AP e RR não fazem mais parte do cluster do Nordeste, que inclui outros estados da região Norte, como AC e AM. Diferentemente dos períodos anteriores, SP, RJ e DF não formam mais um cluster em separado; a região Sul passou a formar um único cluster com o Sudeste. A região Nordeste, entretanto, passa a incluir três estados do Norte (AM, PA, AC). O dendograma para o ano 2000 permite concluir que apenas dois clusters de desenvolvimento são identificados no Brasil: as regiões Sul-Sudeste-Centro-Oeste (SSC) e Norte-Nordeste (NN).

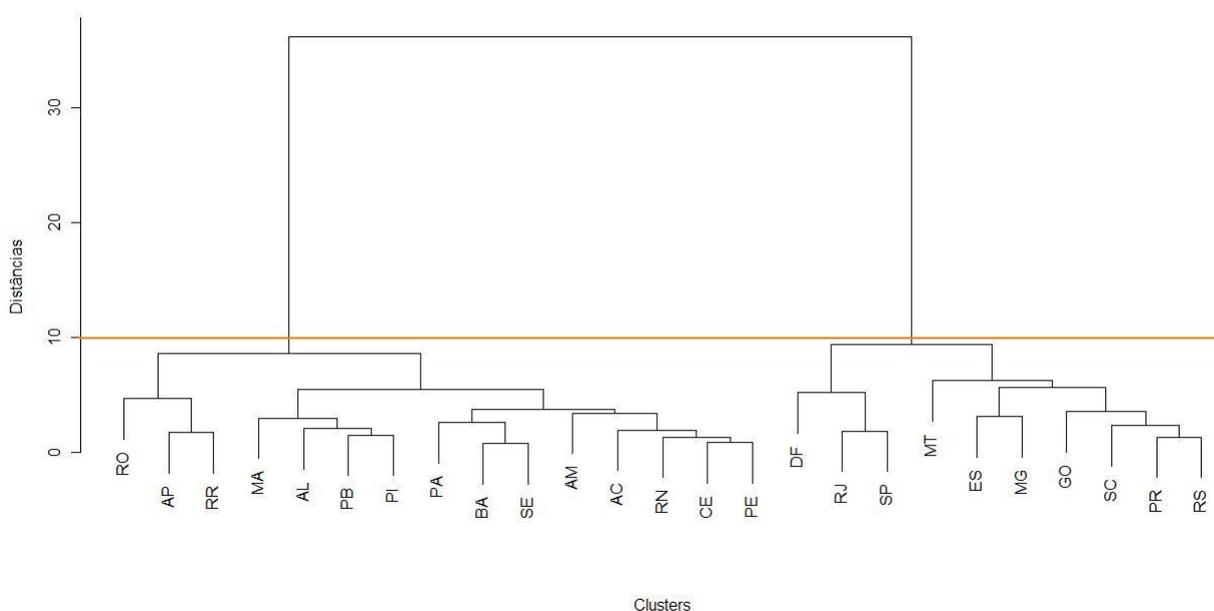


Figura 5: Dendograma mostrando os resultados da aglomeração de unidades da federação, 2000.
Fonte: elaboração própria.

A seguir, nas Tabelas 3 e 4 são apresentados os resultados para as diferenças das distâncias entre os estados brasileiros ao longo dos anos considerados (1970, 1980, 1991 e 2000). Como foi observado anteriormente, foi utilizado o teste não paramétrico de Wilcoxon para dados pareados, uma vez que não depende das suposições de normalidade e homogeneidade das variâncias. Na Tabela 3 aparecem os resultados para o teste da hipótese de igualdade contra a hipótese bicaudal, e, na Tabela 4 os resultados referem-se ao teste de igualdade contra a hipótese unicaudal à esquerda, que indica divergência ao longo dos anos.

Tabela 3: Resultado para a diferença das distâncias entre todos os estados brasileiros

Ano	Mediana
1970	3,91 ^A
1980	4,43 ^B
1991	4,70 ^B
2000	4,77 ^{B*}

Fonte: elaboração própria.

Nota: os valores seguidos pelas mesmas letras não diferem estatisticamente; as letras diferentes indicam que os valores diferem estatisticamente a 0,05 de significância; (*) estatisticamente significativo a 0,10.

Tabela 4: Resultado para a diferença das distâncias entre todos os estados brasileiros

Ano	Mediana
1970	3,91 ^A
1980	4,43 ^B
1991	4,70 ^B
2000	4,77 ^B

Fonte: elaboração própria.

Nota: os valores seguidos pelas mesmas letras não diferem estatisticamente; as letras diferentes indicam que os valores diferem estatisticamente a 0,05 de significância.

Os resultados apresentados acima permitem constatar que: primeiro, as distâncias entre os estados brasileiros, em termos de desenvolvimento econômico, aumentaram ao longo do tempo: esse aumento é estatisticamente significativo. Segundo, como indica a Tabela 4, a hipótese divergência vem ganhando respaldo ao longo do tempo, como referida por Myrdal (1968) e Kaldor (1978), uma vez que a hipótese nula de igualdade foi rejeitada em todos os casos, em favor da hipótese alternativa de divergência no desenvolvimento.

Outra constatação, corroborada pela análise de cluster anteriormente, é que nos anos 1970 a distância entre os estados era menor, quando comparada com a configuração do ano 2000. Embora não haja diferenças significativas entre as distâncias dos anos 1980, 1991 e 2000, o comportamento da mediana parece indicar claramente o seu aumento ao longo do tempo. De 1991, em termos de mediana, as distâncias aumentaram 1,49%. Porém, de 1980 para 1991 esse aumento foi de 6,09%, que é sensivelmente maior que o período posterior.

4. Conclusões

O principal objetivo do trabalho foi realizar uma análise exploratória de uma base de dados para a economia brasileira a nível estadual, levando-se em conta 14 variáveis que expressam 6 dimensões do desenvolvimento econômico, tais como o desenvolvimento industrial, a produtividade agrícola, o capital humano, a pobreza, a desigualdade e as condições de moradia.

A partir dos resultados da análise de Cluster foi possível constatar que o País está diminuindo a heterogeneidade *dentro* de dois grupos distintos de estados, porém, há um aumento das distâncias *entre* esses grupos de estados. Após quatro décadas de desenvolvimento econômico, apenas dois grupos de estados se configuram no Brasil: as macrorregiões Sul-Sudeste-Centro-Oeste e a Norte-Nordeste. Há evidência de que ambas, internamente, estão convergindo em várias dimensões do desenvolvimento, porém, enquanto grupo, ambas divergem entre si. É importante notar que essas diferenças são estatisticamente significativas, configurando um processo sistemático ao longo do tempo, não casual.

Nota-se que, a despeito da relativa similaridade entre os grupos existente nos início da década de 1970, as diferenças acentuaram-se ao longo dos últimos quarenta anos de política econômica (a mediana das distâncias aumentou ao longo do tempo). Esta conclusão deriva do teste para a diferença das distâncias euclidianas, obtidas a partir da matriz de dados padronizada para

cada ano analisado. Isto é, ao longo do tempo, as diferenças entre os grupos foram se acumulando, culminando na formação de apenas dois grupos distintos entre si. Os resultados dos testes para diferença das médias permitem concluir que com relação às condições iniciais, a distância entre os estados brasileiros, tendo por base o ano de 1970, aumentou sistematicamente, apesar desse aumento não ser estatisticamente significativo quando são comparados os anos 1980 com 1991 e 2000.

A estatística Λ de Wilks, empregada para testar a significância das diferenças entre os grupos, indicou a formação sistemática de grupos diferentes ao longo dos anos. Em particular, apesar da migração dos estados de alguns clusters iniciais para outros mais desenvolvidos, as diferenças entre os grupos são estatisticamente significativas, mesmo quando se considera a formação de apenas dois clusters (duas macrorregiões, Sul-Sudeste-Centro-Oeste e Norte-Nordeste).

Um trabalho futuro poderá investigar os principais fatores que influenciaram o desenvolvimento das dimensões estudadas neste trabalho, que, dentro dos limites deste estudo, com caráter exploratório, não puderam ser objeto de análise.

5. Referências bibliográficas

- ARYEETAY, E.; OWUSU, G., MENSAH, E. J. (2009) “An Analysis of Poverty and Regional Inequalities in Ghana”, *GDN Working Paper Series n° 27*, New Delhi.
- BARRO, R.; SALA-I-MARTIN, X. *Economic Growth*, New Delhi: Prentice-Hall.
- CARVALHO, D. F.; SANTANA, A. C.; MENDES, F. A. T. (2006) “Análise de *cluster* da indústria de móveis de madeira do Pará”, *Novos Cadernos NAEA*, vol. **9**(2): 25-54.
- DIXON, R.; THIRLWALL, A. P. (1975) “A model of regional growth-rate differences on kaldorian lines”, *Oxford Economic Papers (New Series)*, vol. **27** (2): 201-214.
- DURLAUF, N. S.; JOHNSON, P. (1995) “Multiple regimes and cross country growth behavior”, *Journal of Applied Econometrics*, vol. **10**(4): pp. 365-384.
- FERREIRA, R. T.; CRUZ, M. S. (2008) “Clubes de convergência na desigualdade de renda nos municípios brasileiros”, In: *Anais do XXXVI Encontro Nacional de Economia (Anpec)*, Salvador, BA.
- GALOR, O. “Convergence? Inference from Theoretical Models”, *Economic Journal*, vol. **106**(437): pp. 1056-1069.
- GLIGOR, M.; AUSLOOS, M. (2011) “Convergence and cluster structures in EU area according to fluctuations in macroeconomic series”, Disponível: <http://arxiv.org/ftp/arxiv/papers/0805/0805.3071.pdf> ; Acesso: 20 Out 2011.
- GRANGER, C. W. (1993) “What we learning about the long run”, *Economic Journal*, **103**: 307-317.
- KALDOR, N. (1978) “The case for regional policies”, In: *Further Essays on Economic Theory*, London: Duckworth, pp. 139-154.
- MARTIN, R.; SUNLEY, P. (1998) “Slow Convergence? New Endogenous Growth Theory and Regional Development”, *Economic Geography*, vol. **74**(3): pp. 201-227.
- MARQUES, A. M. (2009), “Crescimento e grau de desigualdade no Rio Grande do Sul: uma lição a partir dos anos 1990”, *Estudos do Cepe*, n° **30**: pp. 6-29.
- MYRDAL, G. (1968) *Teoria Econômica e Regiões Subdesenvolvidas*, Rio de Janeiro: Saga.

- MOREIRA, R. C.; BRAGA, M. J.; TOYOSHIMA, S. H. (2010) “Crecimiento e desigualdade: prosperidade *versus* armadilhas da pobreza no desenvolvimento econômico dos estados brasileiros”, *Anais do 38º Encontro Nacional de Economia*, Salvador (BA). Disponível: www.anpec.org.br/encontro2010 ; acesso 05 Jan 2012.
- ROMER, D. (2006) *Advanced Macroeconomics*, New York: McGraw-Hill.
- MINGOTI (2005) *Análise de Dados através de Métodos de Estatística Multivariada*. Belo Horizonte: Ed. UFMG.
- JOHNSON, R. A.; WICHERN, D. W. (1998) *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall.
- PEREIRA, J. C. R. (2001) *Análise de dados qualitativos*. São Paulo: Edusp.
- REIS, E. (2001) *Estatística Multivariada Aplicada*. Lisboa: Ed. Silabo.
- R DEVELOPMENT CORE TEAM (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- SIEGEL, S.; CASTELLAN, N. J. (1988) *Nonparametric Statistics for the Behavioral Sciences*, New York: McGraw-Hill.
- SAINT-ARNAUD, S.; BERNARD, P. (2003) “Convergence or Resilience? A Hierarchical Cluster Analysis of the Welfare Regimes in Advanced Economies”, *Current Sociology*, vol. **51**(5): 499-527.
- SETTERFIELD, M. (1997) “‘History versus equilibrium’ and the theory of economic growth”, *Cambridge Journal of Economics*, **21**: 365-78.
- THIRLWALL, A. P. (2005) *A natureza do crescimento econômico*, Brasília: IPEA.

Anexo 2: Matrices de distâncias multidimensionais entre as unidades da federação nos anos de 1970, 1980, 1991 e 2000.

1970

	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP
AC	0.0	2.2	2.10	4.06	3.77	4.2	8.4	3.50	2.43	1.9	4.72	2.04	2.32	4.2	3.72	2.7	2.96	9.3	5.0	3.12	2.42	5.36	4.05	2.7	10.3
AL	2.2	0.0	2.30	4.64	2.74	2.9	8.7	3.16	2.46	2.7	3.93	2.65	2.72	2.7	2.10	1.9	3.43	9.7	3.7	3.90	3.30	5.68	4.70	1.0	10.4
AM	2.1	2.3	0.00	3.22	3.37	3.8	7.2	2.05	0.99	3.4	3.35	1.12	1.19	3.8	2.68	3.0	2.04	8.3	4.7	3.07	1.90	4.09	2.91	2.3	9.3
AP	4.1	4.6	3.22	0.00	5.22	5.8	7.0	3.96	3.47	5.0	4.64	2.87	3.30	5.8	4.75	5.2	3.71	7.0	6.3	3.49	3.27	4.41	3.10	4.4	7.8
BA	3.8	2.7	3.37	5.22	0.00	1.3	8.6	2.90	3.10	4.4	3.91	3.66	3.09	1.7	2.45	3.4	3.86	9.6	2.3	4.57	3.42	5.29	4.75	2.3	9.9
CE	4.2	2.9	3.81	5.84	1.30	0.0	9.3	3.63	3.67	4.6	4.65	4.21	3.53	0.7	2.72	3.1	4.69	10.5	1.7	5.32	4.14	6.23	5.55	2.4	10.9
DF	8.4	8.7	7.23	6.96	8.56	9.3	0.0	6.89	6.77	10.1	6.81	7.60	7.46	9.5	7.68	9.9	6.81	4.3	9.7	6.36	6.71	5.03	6.68	8.4	6.2
ES	3.5	3.2	2.05	3.96	2.90	3.6	6.9	0.00	1.57	4.7	2.07	2.43	2.06	3.8	2.55	3.9	1.77	8.1	4.7	3.88	2.49	3.11	2.43	2.9	8.5
GO	2.4	2.5	0.99	3.47	3.10	3.7	6.8	1.57	0.00	3.8	3.04	1.66	1.21	3.7	2.49	3.3	1.72	8.1	4.5	2.87	1.59	3.60	2.72	2.3	9.0
MA	1.9	2.7	3.38	4.97	4.43	4.6	10.1	4.69	3.80	0.0	5.84	2.99	3.35	4.3	4.65	2.2	4.27	10.7	5.2	4.62	3.87	6.83	5.16	3.2	11.7
MG	4.7	3.9	3.35	4.64	3.91	4.6	6.8	2.07	3.04	5.8	0.00	3.69	3.75	4.7	2.66	5.0	3.07	7.7	5.4	4.53	3.98	3.71	3.64	3.6	7.6
MT	2.0	2.6	1.12	2.87	3.66	4.2	7.6	2.43	1.66	3.0	3.69	0.00	1.38	4.2	3.33	3.1	1.99	8.3	5.1	3.23	2.06	4.21	2.62	2.7	9.2
PA	2.3	2.7	1.19	3.30	3.09	3.5	7.5	2.06	1.21	3.4	3.75	1.38	0.00	3.6	3.09	3.0	2.27	8.6	4.3	3.22	1.44	4.18	2.73	2.5	9.5
PB	4.2	2.7	3.82	5.82	1.68	0.7	9.5	3.79	3.72	4.3	4.70	4.18	3.55	0.0	2.64	2.8	4.83	10.7	1.4	5.31	4.23	6.52	5.69	2.2	11.1
PE	3.7	2.1	2.68	4.75	2.45	2.7	7.7	2.55	2.49	4.6	2.66	3.33	3.09	2.6	0.00	3.4	3.58	8.9	3.3	4.06	3.57	5.02	4.65	1.5	9.3
PI	2.7	1.9	2.95	5.22	3.41	3.1	9.9	3.91	3.31	2.2	5.02	3.08	2.97	2.8	3.40	0.0	4.23	10.9	3.9	5.09	3.95	6.63	5.16	2.1	11.7
PR	3.0	3.4	2.04	3.71	3.86	4.7	6.8	1.77	1.72	4.3	3.07	1.99	2.27	4.8	3.58	4.2	0.00	7.8	5.7	3.44	2.35	2.72	1.83	3.5	8.5
RJ	9.3	9.7	8.34	6.96	9.59	10.5	4.3	8.09	8.08	10.7	7.68	8.32	8.56	10.7	8.92	10.9	7.83	0.0	10.7	7.28	7.69	6.34	7.40	9.4	4.2
RN	5.0	3.7	4.72	6.28	2.32	1.7	9.7	4.66	4.52	5.2	5.40	5.05	4.34	1.4	3.34	3.9	5.72	10.7	0.0	5.56	4.75	7.15	6.46	3.0	11.2
RO	3.1	3.9	3.07	3.49	4.57	5.3	6.4	3.88	2.87	4.6	4.53	3.23	3.22	5.3	4.06	5.1	3.44	7.3	5.6	0.00	2.34	4.47	4.06	3.8	8.5
RR	2.4	3.3	1.90	3.27	3.42	4.1	6.7	2.49	1.59	3.9	3.98	2.06	1.44	4.2	3.57	4.0	2.35	7.7	4.7	2.34	0.00	3.83	2.83	3.1	8.9
RS	5.4	5.7	4.09	4.41	5.29	6.2	5.0	3.11	3.60	6.8	3.71	4.21	4.18	6.5	5.02	6.6	2.72	6.3	7.2	4.47	3.83	0.00	2.44	5.5	6.7
SC	4.1	4.7	2.91	3.10	4.75	5.6	6.7	2.43	2.72	5.2	3.64	2.62	2.73	5.7	4.65	5.2	1.83	7.4	6.5	4.06	2.83	2.44	0.00	4.5	7.9
SE	2.7	1.0	2.26	4.42	2.30	2.4	8.4	2.85	2.28	3.2	3.58	2.68	2.49	2.2	1.54	2.1	3.46	9.4	3.0	3.82	3.13	5.49	4.54	0.0	10.1
SP	10.3	10.4	9.26	7.75	9.87	10.9	6.2	8.48	9.01	11.7	7.56	9.21	9.53	11.1	9.28	11.7	8.51	4.2	11.2	8.49	8.92	6.67	7.95	10.1	0.0

1980

	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP
AC	0.00	2.81	2.91	2.65	3.45	3.7	8.3	6.40	3.63	2.3	4.97	3.00	1.88	3.5	2.98	2.6	4.63	7.78	2.99	3.14	3.11	6.16	5.47	2.29	9.5
AL	2.81	0.00	4.32	4.16	3.25	3.1	9.1	6.63	4.39	3.1	5.03	3.69	3.66	2.3	2.48	2.6	5.16	8.36	2.31	4.16	4.57	6.89	6.50	2.29	9.8
AM	2.91	4.32	0.00	2.35	3.85	4.5	6.6	5.62	2.95	4.8	3.66	3.30	3.38	4.6	3.25	4.6	4.14	5.71	3.69	3.96	2.69	5.08	4.29	3.26	7.5
AP	2.65	4.16	2.35	0.00	4.25	4.9	7.1	5.71	3.32	4.3	4.16	3.68	1.95	4.8	3.67	4.5	4.64	6.25	3.92	3.96	2.94	5.51	4.24	3.43	7.8
BA	3.45	3.25	3.85	4.25	0.00	2.5	7.9	4.35	3.82	4.5	3.61	3.36	2.91	3.0	2.41	3.4	4.25	7.19	2.96	2.63	4.41	6.06	5.92	1.96	8.8
CE	3.69	3.09	4.46	4.88	2.52	0.0	8.4	6.53	4.46	4.5	5.11	4.15	3.79	1.3	1.89	2.4	5.68	8.10	1.66	3.91	5.10	7.46	7.49	1.79	10.3
DF	8.29	9.07	6.60	7.09	7.89	8.4	0.0	7.39	5.63	10.2	5.74	7.03	7.29	8.9	7.24	9.5	6.86	2.64	8.10	7.60	6.00	6.15	6.80	7.77	6.3
ES	6.40	6.63	5.62	5.71	4.35	6.5	7.4	0.00	5.39	7.7	3.76	5.48	5.15	6.9	5.90	7.4	4.91	6.25	6.48	4.63	5.77	5.80	5.36	5.63	6.4
GO	3.63	4.39	2.95	3.32	3.82	4.5	5.6	5.39	0.00	5.6	2.85	1.87	2.99	4.7	2.89	5.1	2.55	5.16	3.94	3.59	1.84	3.64	4.02	3.46	7.4
MA	2.33	3.14	4.84	4.29	4.49	4.5	10.2	7.71	5.63	0.0	6.65	4.73	3.56	4.0	4.36	2.4	6.16	9.62	4.06	4.49	5.16	7.71	6.88	3.58	11.0
MG	4.97	5.03	3.66	4.16	3.61	5.1	5.7	3.76	2.85	6.7	0.00	3.76	4.03	5.5	3.72	6.1	3.39	4.52	4.82	4.36	3.84	4.08	4.06	4.24	6.2
MT	3.00	3.69	3.30	3.68	3.36	4.1	7.0	5.48	1.87	4.7	3.76	0.00	2.75	4.2	2.96	4.4	2.17	6.65	3.79	2.59	2.09	4.10	4.45	3.15	8.4
PA	1.88	3.66	2.38	1.95	2.91	3.8	7.3	5.15	2.99	3.6	4.03	2.75	0.00	3.9	2.94	3.5	3.89	6.64	3.28	2.37	2.71	5.31	4.52	2.38	8.3
PB	3.47	2.26	4.61	4.82	2.97	1.3	8.9	6.87	4.65	4.0	5.47	4.15	3.91	0.0	2.03	2.2	5.79	8.47	1.21	4.14	5.12	7.62	7.54	1.78	10.4
PE	2.98	2.48	3.25	3.67	2.41	1.9	7.2	5.90	2.89	4.4	3.72	2.96	2.94	2.0	0.00	3.0	4.39	6.70	1.45	3.49	3.67	5.97	5.96	1.38	8.8
PI	2.63	2.59	4.57	4.49	3.45	2.4	9.5	7.42	5.12	2.4	6.06	4.44	3.55	2.2	2.99	0.0	6.14	9.12	2.54	4.25	5.21	7.84	7.48	2.34	11.0
PR	4.63	5.16	4.14	4.64	4.25	5.7	6.9	4.91	2.55	6.2	3.39	2.17	3.89	5.8	4.39	6.1	0.00	6.07	5.40	3.59	3.09	2.35	3.25	4.71	7.3
RJ	7.78	8.36	5.71	6.25	7.19	8.1	2.6	6.25	5.16	9.6	4.52	6.65	6.64	8.5	6.70	9.1	6.07	0.00	7.57	7.14	5.70	5.16	5.32	7.24	4.1
RN	2.99	2.31	3.69	3.92	2.96	1.7	8.1	6.48	3.94	4.1	4.82	3.79	3.28	1.2	1.45	2.5	5.40	7.57	0.00	3.90	4.37	7.06	6.81	1.33	9.6
RO	3.14	4.16	3.36	3.96	2.63	3.9	7.6	4.63	3.59	4.5	4.36	2.59	2.37	4.1	3.49	4.3	3.59	7.14	3.90	0.00	3.30	5.45	5.19	2.97	8.6
RR	3.11	4.57	2.69	2.94	4.41	5.1	6.0	5.77	1.84	5.2	3.84	2.09	2.71	5.1	3.67	5.2	3.09	5.70	4.37	3.30	0.00	4.14	3.93	3.79	7.8
RS	6.16	6.89	5.08	5.51	6.06	7.5	6.2	5.80	3.64	7.7	4.08	4.10	5.31	7.6	5.97	7.8	2.35	5.16	7.06	5.45	4.14	0.00	2.50	6.43	6.1
SC	5.47	6.50	4.29	4.24	5.92	7.5	6.8	5.36	4.02	6.9	4.06	4.45	4.52	7.5	5.96	7.5	3.25	5.32	6.81	5.19	3.93	2.50	0.00	6.13	5.5
SE	2.29	2.29	3.26	3.43	1.96	1.8	7.8	5.63	3.46	3.6	4.24	3.15	2.38	1.8	1.38	2.3	4.71	7.24	1.33	2.97	3.79	6.43	6.13	0.00	9.2
SP	9.49	9.79	7.51	7.78	8.76	10.3	6.3	8.40	7.43	11.0	6.15	8.44	8.34	10.4	8.85	11.0	7.29	4.12	9.59	8.58	7.76	6.11	5.54	9.17	0.0

(continua)

Anexo 2: Matrices de distâncias multidimensionais entre as unidades da federação nos anos de 1970, 1980, 1991 e 2000 (continuação).

1991

	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP
AC	0.00	2.94	3.60	3.39	2.89	2.39	7.5	4.33	3.66	4.8	4.64	3.51	3.70	2.2	2.16	2.5	4.28	7.32	1.67	5.10	2.34	5.44	5.77	2.40	9.1
AL	2.94	0.00	4.81	4.92	2.51	2.98	9.1	4.98	5.17	2.7	5.08	5.01	3.26	2.0	3.03	2.6	5.39	8.28	2.38	6.12	4.30	6.67	6.74	2.41	9.8
AM	3.60	4.81	0.00	3.48	4.19	3.85	6.8	3.73	4.11	6.2	4.06	5.18	4.51	4.2	3.32	4.8	3.86	5.72	3.38	6.65	3.80	4.96	5.16	3.19	7.6
AP	3.39	4.92	3.48	0.00	4.88	4.44	6.0	3.53	2.64	6.1	3.96	4.39	4.05	4.5	3.88	5.1	4.17	5.65	3.46	6.33	3.62	5.00	4.77	3.62	7.4
BA	2.89	2.51	4.19	4.88	0.00	1.59	8.3	4.37	5.25	4.2	4.55	5.33	3.12	2.0	1.88	2.3	5.16	7.72	1.93	5.37	3.76	6.44	6.84	1.78	9.6
CE	2.39	2.98	3.85	4.44	1.59	0.00	8.0	4.76	5.02	4.7	4.85	5.39	3.89	1.5	1.37	1.9	5.27	7.57	1.62	6.08	3.63	6.49	6.93	2.08	9.8
DF	7.52	9.05	6.84	5.98	8.32	7.96	0.0	5.65	4.89	10.6	5.50	6.64	7.85	8.6	6.89	9.3	5.26	2.82	7.43	8.97	6.52	4.80	5.72	7.62	5.0
ES	4.33	4.98	3.73	3.53	4.37	4.76	5.7	0.00	2.94	6.1	1.66	4.31	3.21	5.1	3.95	5.7	2.56	4.55	3.91	5.15	4.05	3.45	3.47	3.35	5.7
GO	3.66	5.17	4.11	2.64	5.25	5.02	4.9	2.94	0.00	6.8	3.29	2.60	4.48	5.0	4.04	5.7	2.31	4.67	3.84	5.79	3.43	2.98	3.40	4.11	6.0
MA	4.84	2.73	6.25	6.05	4.15	4.72	10.6	6.14	6.79	0.0	6.29	6.81	3.52	3.9	5.08	3.8	6.94	9.72	4.42	7.21	5.90	8.07	7.58	4.10	11.0
MG	4.64	5.08	4.06	3.96	4.55	4.85	5.5	1.66	3.29	6.3	0.00	4.67	3.39	5.4	3.94	6.0	2.71	4.28	4.24	6.22	4.45	3.39	3.59	3.88	5.7
MT	3.51	5.01	4.19	4.39	5.33	5.39	6.6	4.31	2.60	6.8	4.67	0.00	4.93	5.2	4.49	5.6	3.08	6.63	4.19	4.74	3.27	3.73	4.49	4.64	7.6
PA	3.70	3.26	4.51	4.05	3.12	3.89	7.9	3.21	4.48	3.5	3.39	4.93	0.00	3.9	3.56	4.1	4.39	7.05	3.28	5.34	4.04	5.39	5.03	2.82	8.3
PB	2.21	2.03	4.24	4.48	1.99	1.49	8.6	5.09	5.05	3.9	5.36	5.15	3.93	0.0	2.10	1.4	5.50	8.11	1.38	6.01	3.80	6.81	7.10	1.94	10.1
PE	2.16	3.03	3.32	3.88	1.88	1.37	6.9	3.95	4.04	5.1	3.94	4.49	3.56	2.1	0.00	2.8	4.19	6.47	1.32	5.96	2.94	5.40	6.02	1.96	8.7
PI	2.50	2.56	4.83	5.08	2.35	1.94	9.3	5.68	5.72	3.8	5.97	5.56	4.07	1.4	2.81	0.0	6.05	8.92	2.30	5.95	4.05	7.26	7.45	2.72	10.8
PR	4.28	5.39	3.86	4.17	5.16	5.27	5.3	2.56	2.31	6.9	2.71	3.08	4.39	5.5	4.19	6.0	0.00	4.47	4.32	5.72	3.70	1.47	2.52	4.30	5.4
RJ	7.32	8.28	5.72	5.65	7.72	7.57	2.8	4.55	4.67	9.7	4.28	6.63	7.05	8.1	6.47	8.9	4.47	0.00	6.94	8.90	6.58	4.19	4.82	6.83	3.5
RN	1.67	2.38	3.38	3.46	1.93	1.62	7.4	3.91	3.84	4.4	4.24	4.19	3.28	1.4	1.32	2.3	4.32	6.94	0.00	5.40	3.00	5.63	6.00	1.25	8.9
RO	5.10	6.12	6.65	6.33	5.37	6.08	9.0	5.15	5.79	7.2	6.22	4.74	5.34	6.0	5.86	6.0	5.72	8.90	5.40	0.00	5.08	6.49	6.80	5.15	9.6
RR	2.34	4.30	3.80	3.62	3.76	3.63	6.5	4.05	3.43	5.9	4.45	3.27	4.04	3.8	2.94	4.1	3.70	6.58	3.00	5.08	0.00	4.66	5.30	3.41	8.5
RS	5.44	6.67	4.96	5.00	6.44	6.49	4.8	3.45	2.98	8.1	3.39	3.73	5.39	6.8	5.40	7.3	1.47	4.19	5.63	6.49	4.66	0.00	2.02	5.65	4.7
SC	5.77	6.74	5.16	4.77	6.84	6.93	5.7	3.47	3.40	7.6	3.59	4.49	5.03	7.1	6.02	7.5	2.52	4.82	6.00	6.80	5.30	2.02	0.00	5.81	4.6
SE	2.40	2.41	3.19	3.62	1.78	2.08	7.6	3.35	4.11	4.1	3.88	4.64	2.82	1.9	1.96	2.7	4.30	6.83	1.25	5.15	3.41	5.65	5.81	0.00	8.6
SP	9.12	9.83	7.62	7.39	9.64	9.77	5.0	5.72	6.02	11.0	5.69	7.61	8.25	10.1	8.67	10.8	5.44	3.48	8.87	9.58	8.51	4.72	4.65	8.59	0.0

2000

	AC	AL	AM	AP	BA	CE	DF	ES	GO	MA	MG	MT	PA	PB	PE	PI	PR	RJ	RN	RO	RR	RS	SC	SE	SP
AC	0.00	2.80	3.14	3.16	2.27	1.89	8.3	6.24	4.45	4.0	5.09	5.33	3.07	1.67	1.69	2.24	5.01	7.13	1.51	3.21	2.73	5.91	6.88	1.99	7.79
AL	2.80	0.00	3.71	5.02	2.04	2.35	9.4	7.17	6.12	2.2	5.97	6.56	3.08	1.90	2.57	2.04	6.30	8.35	2.78	5.07	4.94	7.35	8.24	1.93	8.94
AM	3.14	3.71	0.00	3.74	2.77	2.65	8.3	6.33	4.94	4.7	4.77	6.14	3.49	3.26	2.48	3.78	4.93	6.23	2.46	4.66	4.05	5.75	6.71	2.72	7.15
AP	3.16	5.02	3.74	0.00	4.14	3.53	6.9	6.38	3.38	5.9	4.75	5.37	4.24	3.99	3.02	4.79	4.63	5.71	2.88	4.48	1.77	5.48	6.50	3.75	6.60
BA	2.27	2.04	2.77	4.14	0.00	1.88	8.5	5.49	5.07	2.9	4.58	5.94	2.08	1.73	1.92	2.37	4.98	7.18	1.90	3.47	3.87	6.00	6.92	0.80	7.75
CE	1.89	2.35	2.65	3.53	1.88	0.00	8.1	6.43	4.91	3.7	5.05	6.03	2.68	2.05	0.89	2.09	5.30	7.17	1.50	4.19	3.70	6.23	7.21	1.79	7.93
DF	8.29	9.43	8.31	6.91	8.47	8.14	0.0	7.42	5.25	10.5	6.24	7.56	8.53	8.87	7.55	9.59	5.49	4.17	7.48	8.17	6.87	5.49	6.08	8.36	4.66
ES	6.24	7.17	6.33	6.38	5.49	6.43	7.4	0.00	5.32	7.6	3.18	6.70	5.24	6.59	6.01	7.46	4.18	5.94	5.77	4.16	5.78	4.55	4.90	5.71	5.60
GO	4.45	6.12	4.94	3.38	5.07	4.91	5.3	5.32	0.00	7.1	3.28	3.62	4.98	5.29	4.31	6.28	2.08	3.98	3.97	4.70	2.88	2.88	4.03	4.94	4.31
MA	3.95	2.16	4.72	5.95	2.86	3.73	10.5	7.56	7.12	0.0	6.55	7.70	3.28	2.78	4.04	2.91	7.13	9.25	3.92	5.45	5.69	8.07	8.72	2.76	9.66
MG	5.09	5.97	4.77	4.75	4.58	5.05	6.2	3.18	3.28	6.6	0.00	5.34	3.84	5.55	4.55	6.54	2.17	4.16	4.24	4.53	4.38	2.81	3.36	4.65	3.94
MT	5.33	6.56	6.14	5.37	5.94	6.03	7.6	6.70	3.62	7.7	5.34	0.00	6.01	6.28	5.66	6.93	3.90	6.63	5.55	5.68	4.76	4.60	5.82	6.06	6.71
PA	3.07	3.08	3.49	4.24	2.08	2.68	8.5	5.24	4.98	3.3	3.84	6.01	0.00	3.14	2.73	3.67	4.75	7.09	2.58	3.88	4.11	5.59	6.26	2.22	7.34
PB	1.67	1.90	3.26	3.99	1.73	2.05	8.9	6.59	5.29	2.8	5.55	6.28	3.14	0.00	2.05	1.55	5.70	7.66	1.87	3.86	3.63	6.71	7.60	1.28	8.35
PE	1.69	2.57	2.48	3.02	1.92	0.89	7.5	6.01	4.31	4.0	4.55	5.66	2.73	2.05	0.00	2.57	4.76	6.49	1.00	3.95	3.16	5.74	6.75	1.66	7.28
PI	2.24	2.04	3.78	4.79	2.37	2.09	9.6	7.46	6.28	2.9	6.54	6.93	3.67	1.55	2.57	0.00	6.66	8.68	2.75	4.45	4.56	7.57	8.48	2.23	9.41
PR	5.01	6.30	4.93	4.63	4.98	5.30	5.5	4.18	2.08	7.1	2.17	3.90	4.75	5.70	4.76	6.66	0.00	3.62	4.39	4.52	3.93	1.36	2.67	5.07	3.57
RJ	7.13	8.35	6.23	5.71	7.18	7.17	4.2	5.94	3.98	9.2	4.16	6.63	7.09	7.66	6.49	8.68	3.62	0.00	6.19	6.97	5.58	3.51	3.97	7.03	1.89
RN	1.51	2.78	2.46	2.88	1.90	1.50	7.5	5.77	3.97	3.9	4.24	5.55	2.58	1.87	1.00	2.75	4.39	6.19	0.00	3.59	2.77	5.32	6.23	1.58	6.90
RO	3.21	5.07	4.66	4.48	3.47	4.19	8.2	4.16	4.70	5.5	4.53	5.68	3.88	3.86	3.95	4.45	4.52	6.97	3.59	0.00	3.45	5.13	5.90	3.58	7.26
RR	2.73	4.94	4.05	1.77	3.87	3.70	6.9	5.78	2.88	5.7	4.38	4.76	4.11	3.63	3.16	4.56	3.93	5.58	2.77	3.45	0.00	4.76	5.77	3.56	6.31
RS	5.91	7.35	5.75	5.48	6.00	6.23	5.5	4.55	2.88	8.1	2.81	4.60	5.59	6.71	5.74	7.57	1.36	3.51	5.32	5.13	4.76	0.00	1.53	6.10	3.12
SC	6.88	8.24	6.71	6.50	6.92	7.21	6.1	4.90	4.03	8.7	3.36	5.82	6.26	7.60	6.75	8.48	2.67	3.97	6.23	5.90	5.77	1.53	0.00	6.99	3.01
SE	1.99	1.93	2.72	3.75	0.80	1.79	8.4	5.71	4.94	2.8	4.65	6.06	2.22	1.28	1.66	2.23	5.07	7.03	1.58	3.58	3.56	6.10	6.99	0.00	7.65
SP	7.79	8.94	7.15	6.60	7.75	7.93	4.7	5.60	4.31	9.7	3.94	6.71	7.34	8.35	7.28	9.41	3.57	1.89	6.90	7.26	6.31	3.12	3.01	7.65	0.00

Anexo 2: Médias das variáveis e número de unidades da federação por cluster e por período.

1970												
Clusters	krrpc	krupe	cerpc	ceipc	theil	p0	idhl	idhe	idhr	yappc	yatpc	er
SUD	16.471	83.087	0.2102	0.3065	0.4966	37.897	0.47300	0.6356	0.8430	0.0297	13.266	0
SUL	11.182	43.754	0.0583	0.0993	0.4783	73.810	0.4476	0.5042	0.3450	0.0368	0.6853	0
NE	0.6468	29.446	0.0268	0.0405	0.5200	87.898	0.3823	0.3265	0.1899	0.1039	0.5775	0
TODOS	0.9931	42.751	0.0639	0.1007	0.4972	75.136	0.4245	0.4489	0.3427	0.0628	0.7191	0
1980												
Clusters	krrpc	krupe	cerpc	ceipc	theil	p0	idhl	idhe	idhr	yappc	yatpc	er
SUD	22.320	101.359	0.5715	0.7761	0.5800	16.363	0.5493	0.6990	0.9626	0.0685	0.1348	0
SUL	15.416	64.033	0.1805	0.7382	0.5760	35.334	0.5716	0.6210	0.9226	0.1507	0.6758	0
NE	0.6454	36.776	0.0890	0.1887	0.6150	69.473	0.4672	0.3940	0.4378	0.0691	0.2560	0
NO	0.9007	39.213	0.1833	0.1672	0.5688	48.006	0.5274	0.5196	0.7327	0.0841	0.3513	0
TODOS	11.069	50.855	0.1992	0.3613	0.5864	48.544	0.5196	0.5212	0.7039	0.0908	0.3598	0
1991												
Clusters	krrpc	krupe	cerpc	ceipc	theil	p0	idhl	idhe	idhr	yappc	yatpc	er
SUD	20.828	93.878	0.5519	0.6334	0.6420	18.349	0.7170	0.8460	0.7660	0.0779	0.1392	0
SUL	11.860	57.295	0.3303	0.6015	0.6320	35.536	0.6891	0.7761	0.6707	0.1258	0.4982	0
NE	0.6546	36.564	0.1677	0.4779	0.7334	65.969	0.5886	0.6112	0.5616	0.0820	0.2447	0
NO	0.5322	29.639	0.2632	0.1327	0.6872	44.939	0.6438	0.7122	0.6420	0.1839	0.3348	0
TODOS	0.9503	47.861	0.2784	0.4621	0.6848	47.527	0.6432	0.7057	0.6327	0.1141	0.3211	0
2000												
Clusters	krrpc	krupe	cerpc	ceipc	theil	p0	idhl	idhe	idhr	yappc	yatpc	er
SUL-SUD	30.980	98.055	0.5683	0.7436	0.6527	22.153	0.7574	0.8858	0.7516	0.0859	0.3712	0
NO-NE	0.9309	45.305	0.3037	0.3826	0.7481	52.074	0.6784	0.7831	0.6281	0.0452	0.1093	0
TODOS	17.978	66.405	0.4095	0.5270	0.7099	40.105	0.7100	0.8242	0.6775	0.0615	0.2141	0

Fonte: elaboração própria.