

EL DESARROLLO SOCIO-ECONOMICO DE LAS PROVINCIAS CHILENAS: UNA CLASIFICACION BASADA EN EL ANALISIS DE COMPONENTES PRINCIPALES

Prof. ROBIN M. HAYNES

ABSTRACT

The level of socio-economic development is measured in a single index for each province of Chile by condensing the information of thirty-one variables using principal components analysis. The provinces are then classified according to their development level and population. The classification method consists of successively grouping the most similar provinces together. Four main groups of provinces are recognized: (1) Santiago, (2) Valparaíso and Concepción, (3) the northern and extreme southern provinces and (4) the rural provinces of central and southern Chile. The classification method can be used with more than two variables and it will also group small areas into regions, if required.

RÉSUMÉ

Le niveau de développement social et économique, peut se mesurer avec un seul indice pour chaque province du Chili, basé sur l'information apportée par trente et une variables, utilisant l'analyse des composants principaux.

Les provinces se classifient, alors, selon leurs niveaux de développement et leur population. La méthode de classification se réalise avec la condensation des principales provinces les plus semblables. On reconnaît quatre groupes principaux dans ces provinces:

(1) Santiago, (2) Valparaíso et Concepción, (3) les provinces de Nord et les plus méridionales, et (4) les provinces rurales du centre et du sud du Chili. La méthode de classification, peut être utilisée avec plus de deux variables et aussi réunir de petites zones pour former des régions si on le désire ainsi.

RESUMEN

El nivel de desarrollo socio-económico se mide en un solo índice para cada provincia de Chile, condensando la información de treinta y una variables, usando el análisis de componentes principales. Las provincias se clasifican entonces según su nivel de desarrollo y población. El método de clasificación consiste en la agrupación sucesiva de las provincias más similares. Se reconocen cuatro grupos principales de provincias: (1) Santiago, (2) Valparaíso y Concepción, (3) las provincias nortinas y del extremo sur y (4) las provincias rurales del centro y sur de Chile. El método de clasificación puede usarse con más de dos variables y también agrupará pequeñas áreas para formar regiones, si se desea.

INTRODUCCION

Si bien cada una de las veinticinco provincias de Chile tiene características propias, algunos grupos de provincias tienen características similares. La tarea de planificación nacional se hace más fácil clasificando las provincias en diferentes grupos, cada grupo con sus circunstancias y problemas que le son particulares. Naturalmente, existen muchos criterios que se podrían usar para hacer tal clasificación. Dos de los más bá-

sicos son la población de la provincia y el nivel de desarrollo socio-económico de la provincia, ya que estos dos factores deberían influir en estrategias de planificación futura. Estos son los dos criterios usados aquí para clasificar las provincias de Chile, usando un método objetivo que sucesivamente agrupa las provincias similares. Primero, sin embargo, es necesario describir cómo se ha medido el desarrollo.

MEDIDA DE DESARROLLO

"Desarrollo socio-económico" implica mezclar varias características diferentes, pero relacionadas: El desarrollo no se puede medir en un censo, porque representa la síntesis de muchas medidas. Las 31 variables consideradas por Mattelart y Garretón (1), por ejemplo, están todas relacionadas por el concepto de desarrollo. Pero examinar simultáneamente la variación de 31 características diferentes a través de todo Chile es una tarea muy difícil, porque la red de relaciones recíprocas es tan complicada. Lo que se necesita es un método para condensar la información proporcionada por las medidas de las 31 variables. Hay varias maneras de hacer esto, incluso aquella adoptada por Mattelart y Garretón, pero el método que tienden a favorecer los estadísticos que tienen acceso a una computadora es el "análisis de componentes principales". Esta es la técnica que se usará aquí con los datos colectados por Mattelart.

Las variables se refieren a urbanización, industrialización, salud, estándar de vida y niveles culturales. Son las siguientes:

1. Producto bruto interno a.p.m. anual per cápita EO (País = 100).
2. Consumo energía eléctrica KWH. Neto. Per cápita anual.
3. Porcentaje de población en ciudades de más de 20.000 habitantes.
4. Porcentajes estratos medios urbanos en población activa masculina.
5. Porcentaje población activa agrícola (hombres y mujeres).
6. Porcentaje población activa: sector secundario (hombres y mujeres).
7. Porcentaje población activa: sector terciario (hombres y mujeres).
8. Porcentaje de menores activos (sexo masculino).
9. Porcentaje de personas de 65-84 años activas (sexo masculino).
10. Porcentaje de mujeres activas en actividades no agrícolas.
11. Porcentaje población activa sindicalizada.
12. Tasa mortalidad infantil por 1.000 nacimientos.
13. Tasa mortalidad neumonía por 100.000 habitantes.
14. Porcentaje certificación médica de las defunciones.
15. Número de camas hospitalarias por mil habitantes.
16. Número de médicos hospitalarios por 100.000 habitantes.
17. Número de matronas por 100.000 mujeres (15-49 años).
18. Número de enfermeras y practicantes por 100.000 habitantes.
19. Porcentaje viviendas con agua potable - cañería.
20. Porcentaje viviendas con alumbrado eléctrico,
21. Porcentaje viviendas sin baño.
22. Ahorro - vivienda per cápita E° al año.
23. Edificación per cápita E° al año.
24. Número de autos particulares por mil habitantes.
25. Porcentaje analfabetos funcionales de más de 15 años.
26. Porcentaje población con instrucción secundaria.

(1) Mattelart, Armand y Garretón, Manuel: *"Integración Nacional y Marginalidad"*. Editorial del Pacífico S.A. Santiago, Chile. 1965. pp. 38 - 39.

27. Porcentaje población con instrucción universitaria.
28. Coeficiente atención primaria.
29. Porcentaje asistentes en los matriculados primarios.
30. Porcentaje matrícula media (13-18 años).
31. Número de diarios por 100 habitantes.

Cada variable se midió en las 25 provincias de Chile en 1960.

CORRELACIONES ENTRE LAS VARIABLES

La Tabla I es la matriz de correlación que da los coeficientes de correlación entre cada par de variables. El coeficiente de correlación es una medida muy útil de la fuerza de asociación entre variables. Si dos variables están fuertemente relacionadas entre sí, de tal manera que sea posible predecir el valor de una a partir de la otra, entonces su correlación es alta. Por otra parte, se dice que las variables que están débilmente relacionadas tienen correlaciones bajas. Existe correlación negativa si los valores altos de una variable están asociados con valores bajos de la otra. El coeficiente de correlación expresa la correlación entre las variables en un solo número, que fluctúa entre 0 (ninguna correlación) y 1 (correlación perfecta). Si las variables están negativamente correlacionadas, el coeficiente tiene un signo negativo. Leyendo la primera columna de la Tabla 1, por ejemplo, vemos que la variable 1 tiene una correlación de 0,53 con la variable 2, una correlación de 0,77 con la variable 3, y así sucesivamente.

El hecho de que la mayoría de los coeficientes en la Tabla no tengan signo negativo, demuestra que la mayoría de las variables están relacionadas en forma positiva: al aumentar una, la otra aumenta. Sin embargo, las relaciones que comprenden las variables 5, 8, 9, 12, 13, 21 ó 25 con otras variables son negativas, indicando que las provincias que tienen valores altos en las otras variables tienden a tener valores bajos en estas siete variables. Al identificar estas siete en la lista de variables, encontramos que están todas asociadas con una falta de desarrollo más que-con-desarrollo, al contrario de las otras variables en el estudio.

Si bien hay algunos coeficientes muy bajos, los coeficientes de correlación son generalmente altos, lo que significa que la mayoría de las variables están fuertemente asociadas entre sí. Las agrupaciones de variables que están fuertemente relacionadas se pueden encontrar marcando los coeficientes altos en la tabla (por ejemplo, aquellos mayores a 0,8) y anotando cuáles son las variables que los producen. Las variables 3, 4, 5, 7, 16, 18, 19, 20, 21, 25, 26 y 28 forman una agrupación especialmente fuerte. Si se conociera el valor de sólo una de estas variables para una provincia chilena, sería posible predecir los valores de todas las otras con bastante precisión. Tomadas en conjunto, proporcionan una gran cantidad de información sobre desarrollo socio-económico. Por ejemplo, una provincia con una alta proporción de población urbana también tendrá altas proporciones de estratos medios urbanos, trabajadores terciarios, doctores y practicantes, red de agua potable, electricidad doméstica y personas que por lo menos tienen educación secundaria, pero bajas proporciones de trabajadores agrícolas, viviendas sin baños y analfabetos. Esta colección de características claramente describe una provincia más desarrollada. Una provincia menos desarrollada, por el contrario, es una en que se encuentran proporciones relativamente altas de trabajadores agrícolas, viviendas sin baños y analfabetos, conjuntamente con proporciones relativamente bajas de población urbana, clase media, trabajadores terciarios y servicios domésticos y escasos recursos educacionales.

TABLA I

MATRIZ DE CORRELACION

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	.53														
3	.77	.47													
4	.72	.34	.84												
5	-.86	-.50	-.84	-.88											
6	.56	.23	.82	.77	-.75										
7	.80	.32	.87	.92	-.89	.81									
8	-.49	-.39	-.50	-.63	.68	-.50	-.59								
9	-.41	-.40	-.63	-.54	.52	-.48	-.51	.22							
10	.62	.13	.80	.81	-.71	.88	.85	-.34	-.51						
11	.76	.78	.59	.53	-.80	.40	.54	-.59	-.44	.32					
12	-.67	-.25	-.53	-.69	.57	-.33	-.70	.41	.31	-.47	-.31				
13	-.64	-.29	-.54	-.65	.55	-.33	-.64	.36	.29	-.49	-.39	.83			
14	.72	.41	.66	.69	-.72	.45	.71	-.31	-.65	.55	.57	-.70	-.60		
15	.65	.35	.66	.71	-.67	.55	.76	-.48	-.61	.65	.45	-.71	-.71	.71	
16	.70	.24	.83	.85	-.73	.80	.85	-.48	-.65	.90	.42	-.58	-.57	.67	.75
17	.74	.51	.51	.54	-.60	.29	.55	-.34	-.57	.41	.55	-.71	-.70	.73	.78
18	.78	.59	.82	.81	-.81	.67	.86	-.54	-.62	.73	.56	-.70	-.64	.82	.85
19	.79	.48	.84	.86	-.87	.72	.90	-.57	-.70	.80	.68	-.61	-.61	.79	.81
20	.82	.51	.87	.84	-.85	.71	.84	-.47	-.79	.77	.67	-.62	-.61	.78	.81
21	-.69	-.41	-.86	-.92	.78	-.73	-.86	.54	.71	-.80	-.52	.65	.68	-.75	-.82
22	.70	.51	.76	.87	-.80	.67	.78	-.56	-.51	.78	.58	-.64	-.59	.69	.72
23	.69	.03	.74	.79	-.73	.71	.81	-.34	-.49	.80	.37	-.57	-.51	.75	.56
24	.59	.14	.66	.64	-.50	.61	.70	-.26	-.63	.80	.21	-.56	-.53	.57	.78
25	-.89	-.50	-.90	-.92	.94	-.76	-.94	.61	.60	-.76	-.69	.71	.68	-.78	-.80
26	.80	.28	.87	.90	-.82	.80	.95	-.50	-.61	.90	.47	-.68	-.63	.71	.80
27	.62	.26	.77	.84	-.72	.86	.81	-.45	-.55	.91	.42	-.51	-.46	.60	.61
28	.84	.51	.88	.84	-.88	.66	.89	-.62	-.63	.65	.68	-.70	-.65	.78	.77
29	.39	.27	.54	.49	-.69	.60	.51	-.62	-.27	.39	.55	-.13	-.23	.29	.34
30	.76	.38	.85	.84	-.78	.62	.87	-.56	-.57	.68	.52	-.67	-.69	.68	.77
31	.86	.39	.92	.77	-.78	.75	.84	-.46	-.54	.82	.56	-.55	-.53	.64	.64
17		18	19	20	21	22	23	24	25	26	27	28	29	30	
18	.71														
19	.70	.91													
20	.77	.83	.93												
21	-.65	-.85	-.90	-.91											
22	.57	.78	.84	.80	-.83										
23	.45	.73	.76	.75	-.77	.65									
24	.60	.72	.72	.78	-.78	.65	.67								
25	-.72	-.89	-.92	-.93	-.90	-.82	-.78	-.68							
26	.63	.87	.91	.90	-.91	.90	.84	.85	-.93						
27	.42	.73	.83	.79	-.84	.82	.81	.74	-.77	.87					
28	.72	.91	.91	.89	-.86	.74	.74	.62	-.95	.87	.68				
29	.25	.38	.48	.47	-.40	.37	.35	.09	-.55	.39	.33	.52			
30	.64	.85	.86	.82	-.86	.72	.70	.64	-.89	.86	.65	.94	.41		
31	.57	.81	.83	.85	-.79	.74	.78	.76	-.86	.90	.79	.84	.38	.83	

UNA MEDIDA COMPUESTA DE DESARROLLO

El análisis de componentes principales es un método estadístico usado para fabricar artificialmente nuevas variables compuestas a partir de la materia prima de los datos originales. El método condensa la información contenida en muchas variables con relación recíproca en un pequeño número de variables compuestas que no están relacionadas entre sí. Aquí nos interesa combinar 31 variables para crear una medida única de desarrollo socio-económico, así es que el objeto es definir una sola medida compuesta. No se describirá exactamente como se hace esto: requiere operaciones de álgebra de matrices efectuadas por computadora. El resultado de este proceso es que se encuentra la combinación óptima de las 31 variables, de tal manera que la nueva medida compuesta tiene la mayor correlación posible con todas las variables.

La medida compuesta que expresa la mayor información posible a partir de las 31 variables originales se llama un "componente principal". Los coeficientes de correlación entre las variables y el componente principal se llaman "saturaciones" (loadings). La Tabla II enumera las saturaciones para todas las variables que se han incluido en el estudio.

TABLA II

LAS SATURACIONES

1	0,856
2	0,459
3	0,911
4	0,924
5	-0,907
6	0,783
7	0,943
8	-0,593
9	-0,665
10	0,838
11	0,647
12	-0,712
13	-0,692
14	0,803
15	0,836
16	0,909
17	0,719
18	0,924
19	0,958
20	0,951
21	-0,941
22	0,865
23	0,818
24	0,771
25	-0,977
26	0,959
27	0,848
28	0,939
29	0,498
30	0,895
31	0,894

Vemos, a partir de la tabla, que la medida compuesta está bien correlacionada con la mayor parte de las variables originales, siendo las únicas excepciones sobresalientes las bajas correlaciones con las variables 2, 8 y 29. En la tabla hay tanto altas correlaciones positivas como altas correlaciones negativas. Las correlaciones positivas son con variables que aumentan en valor a medida que aumenta el nivel de desarrollo, y las correlaciones negativas son con variables cuyos valores disminuyen a medida que aumenta el desarrollo (las variables 5, 8, 9, 12, 13, 21 y 25), lo que confirma que el componente principal es una medida de desarrollo.

A pesar de que se ha definido el componente principal como el de mayor similitud a las variables originales, ninguna sola variable se correlaciona perfectamente con él. En efecto, el componente representa exactamente 69,14 % de toda la información contenida en la matriz de los datos originales. Al reducir 31 variables a una sola, se ha perdido 30,86% de la información. Sin embargo, este es un pequeño precio por la enorme simplificación que se logra. Además el 69% que queda en la nueva variable describe la tendencia general más importante de los datos, mientras que la información perdida consiste en variaciones poco comunes.

Las variables 2, 8 y 29, por ejemplo, contribuyen muy poco al componente principal. La razón es que no se puede predecir ninguna de estas tres variables a partir de todas las otras variables en el estudio. Esto se puede confirmar inspeccionando la matriz de correlación.

DESARROLLO Y POBLACION

Si se considera el componente principal como una variable compuesta, entonces debe ser posible medirlo en cada provincia. Esto se hace convirtiendo todos los datos originales a puntajes estándar para eliminar las diferentes unidades de medición y darles a todas un promedio de cero y una desviación estándar de uno. Entonces se suman los 31 puntajes estándar en cada provincia, pero sólo después que cada uno se ha multiplicado por la saturación de la variable apropiada. Esta multiplicación asegura que variables que tienen altas correlaciones con el componente tengan mucho más importancia que las variables con saturaciones bajas. El resultado es un solo valor para cada provincia, que se llama una "calificación de componente" (score).

Las calificaciones de componentes de desarrollo para las provincias aparecen en la Tabla III, en orden decreciente de desarrollo. Más altas en desarrollo están las provincias urbanizadas de Santiago, Valparaíso y Concepción, las provincias del norte y del extremo sur del país. Las calificaciones de desarrollo más bajas ocurren en las provincias rurales del centro y del sur de Chile.

Las provincias de Chile se podrían agrupar según su nivel de desarrollo solamente, pero tal vez sería más útil considerar el desarrollo conjuntamente con la población de la provincia. Esto se debe a que provincias importantes con numerosa población, que tienen escaso desarrollo, constituyen problemas de planificación más serios que aquellas provincias con el mismo desarrollo, pero escasa población. Así, los datos de población de 1960 para las provincias también aparecen en la Tabla III, junto con los valores que resultan al convertir éstos a puntajes estándar. Los puntajes estándar se calculan restando la población media (205.753, omitiendo el valor extremo de Santiago) de la población de cada provincia y dividiendo por la desviación estándar (136.198). Es necesario hacer esto, para poder comparar los valores de desarrollo y de población directamente, y este proceso produce cifras de población del mismo orden de magnitud que las cifras de desarrollo. Para ambas medidas, un valor de cero representa aproximadamente la posición media.

La figura 1 indica las posiciones relativas de las provincias en términos de nivel de desarrollo y población. La provincia de Santiago no está indicada en su posición correcta; debido a su alta población se ubica a una considerable distancia fuera del gráfico.

Los ejes del gráfico dividen las provincias en cuatro diferentes categorías: aquellas con desarrollo y población sobre la media, aquellas con alto desarrollo y baja población, aquellas con desarrollo y población bajo la media y, finalmente, aquellas con desarrollo relativamente bajo, pero poblaciones relativamente altas.

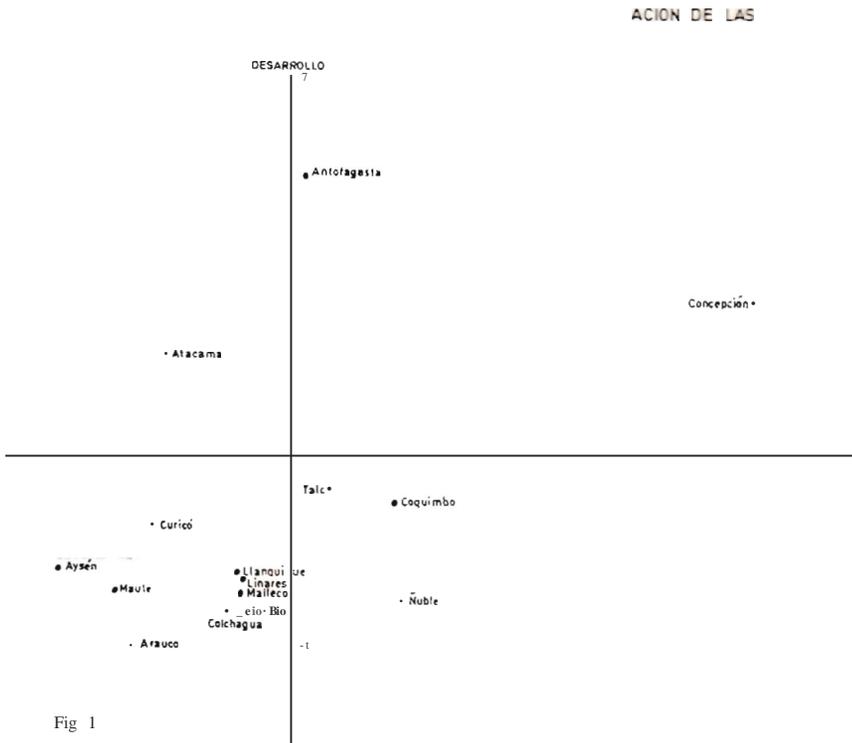


Fig 1

EL PROCESO DE AGRUPAMIENTO

La distancia entre las provincias en el gráfico es una medida de su similitud cuando se consideran tanto el desarrollo como la población. Las dos provincias más cercanas, y por lo tanto el par más similar, solamente en cuanto a estas dos características, son Linares y Llanquihue. Las dos más distantes, y por lo tanto menos parecidas, son Santiago y Aisén.

Si las provincias se agrupan a base de su cercanía en el gráfico, el resultado será grupos de provincias similares. La manera lógica de efectuar tal agrupación es primero unir las dos provincias más similares, después agrupar el segundo par más relacionado, luego el tercer par, y así sucesivamente hasta que al fin todas las provincias estén incluidas en un gran grupo. El investigador o puede entonces seleccionar el número de grupos que le conviene más.

TABLA III

DATOS DE DESARROLLO Y POBLACION

Santiago	2,264	2.436.034	16.375
Valparaíso	1,999	613.164	2.991
Antofagasta	1,494	215.376	0.070
Magallanes	1,619	73.224	-0.973
Tarapacá	1,316	123.064	-0.607
Concepción	0,810	539.440	2.450
Atacama	0,521	116.309	-0.657
Aconcagua	0,351	140.548	-0.479
O'Higgins	0,003	259.724	0.396
Talca	-0,189	206.255	0.004
Osorno	-0,241	144.088	-0.453
Coquimbo	-0,246	280.617	0.550
Curicó	-0,368	105.844	-0.733
Valdivia	-0,487	259.798	0.397
Aisén	-0,589	37.803	-1.233
Llanquihue	-0,617	167.499	-0.281
Linares	-0,649	171.302	-0.253
Maule	-0,703	79.763	-0.925
Malleco	-0,717	170.205	-0.261
Nuble	-0,766	285.703	0.587
Cautín	-0,805	389.864	1.352
Colchagua	-0,813	158.543	-0.347
Bio-Bio	-0,815	168.837	-0.271
Arauco	-1,029	89.504	-0.953
Chiloé	-1,340	182.905	-0.168

Cuando se agrupan dos provincias, sus dos puntos en un gráfico se reemplazan por un solo punto que indica el centro de gravedad del nuevo grupo. La posición del centro de gravedad se define por las medias de los dos valores de desarrollo y de población. Si una tercera provincia está más cerca a este centro de gravedad que a cualquier otra provincia o grupo de provincias, se unirá a las dos, para formar un grupo de tres. El centro de gravedad del nuevo grupo se calculará entonces como los valores medios de población y desarrollo para las tres provincias y se usará de ahí en adelante para representar la posición del nuevo grupo. La observación de estas reglas asegura que la clasificación se hará sólo en base a los datos numéricos, eliminando totalmente la influencia de cualquier idea preconcebida que pudiera tener el investigador.

Las provincias más cercanas, y las primeras en agruparse, son Linares y Llanquihue. El grupo Linares - Llanquihue está representado por un valor de desarrollo de -0,633 y una cifra de población de -0,269 (estas son las medias para las dos provincias). Las segundas provincias más cercanas son Colchagua y Bío-Bío, así es que éstas se agrupan. Su centro de gravedad es -0,814 para desarrollo y -0,309 para población. Los terceros puntos más cercanos son ahora la provincia de Malleco y el centro de gravedad del grupo Linares - Llanquihue, así es que Malleco se une a Linares y Llanquihue, y el centro de gravedad del grupo se corre a -0,611 (desarrollo) y -0,265 (población). La cuarta etapa en el proceso de agrupamiento es para que Colchagua y Bío-Bío se unan a Malleco, Linares y Llanquihue. Después de esto, Atacama y Aconcagua se juntan y entonces Curicó se une a Osorno. Este proceso continúa, con grupos formándose y

uniéndose a otros grupos, por un total de 24 pasos, etapa en la cual todos los grupos se han unido.

Los distintos pasos en el proceso se ven en la figura 2. A partir de este diagrama, la situación en cualquiera etapa se puede evaluar.

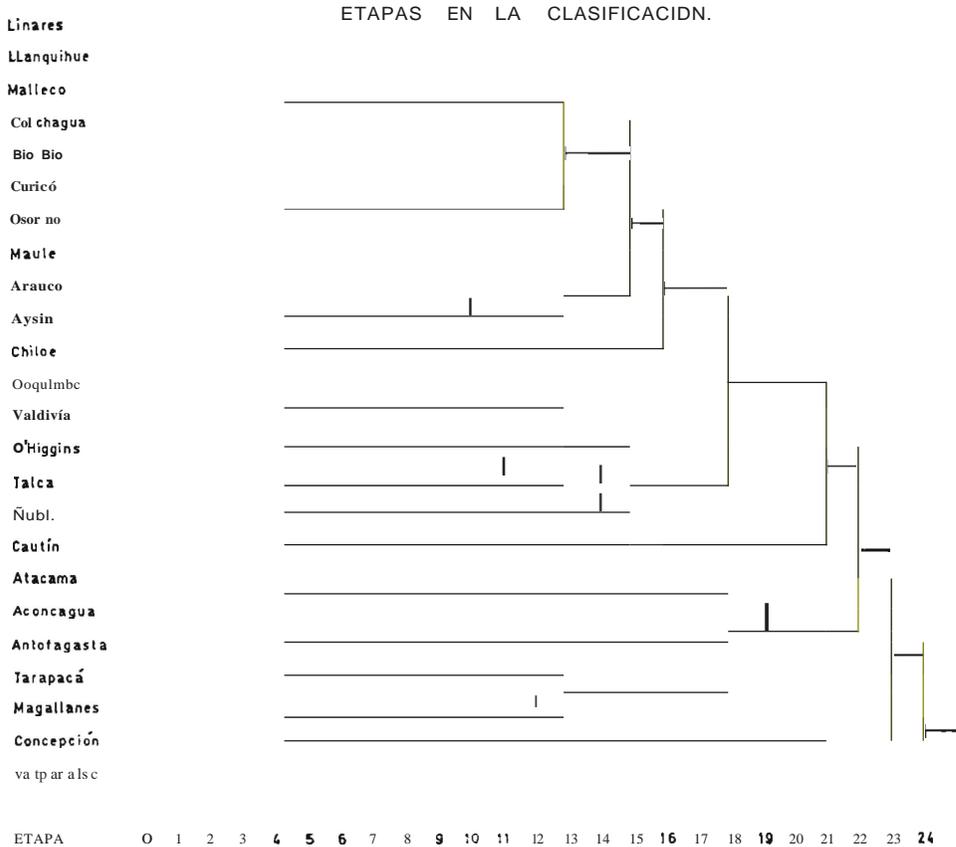


Fig.2

Por ejemplo, después del paso 17, las provincias formaron ocho grupos cuyos integrantes se pueden leer en el diagrama. Estos son los mejores ocho grupos que se pueden formar a partir de las 25 provincias con los datos disponibles. Es decir, si necesitamos ocho grupos, estos grupos dan la semejanza máxima dentro de cada grupo conjuntamente con las diferencias máximas entre grupos. En forma similar, el paso 18 da los mejores siete grupos, y así sucesivamente.

LA CLASIFICACION

La figura 3 muestra la situación hasta e incluyendo el paso 21, quedando cuatro grupos de provincias. El primero consiste en Santiago solo. Con su alto valor de desarrollo y población excepcionalmente alta, hay que considerar esta provincia aparte. El segundo grupo contiene las provincias urbanizadas de Valparaíso y Concepción, que también tienen cifras relativamente altas de población y desarrollo. El tercer grupo es alto en desarrollo, pero relativamente bajo en población. Consiste en las restantes provincias urbanizadas del norte y del extremo sur: Tarapacá, Antofagasta, Atacama, Aconcagua y Magallanes, El cuarto y más grande de

AGRUPACION DE PROVINCIAS

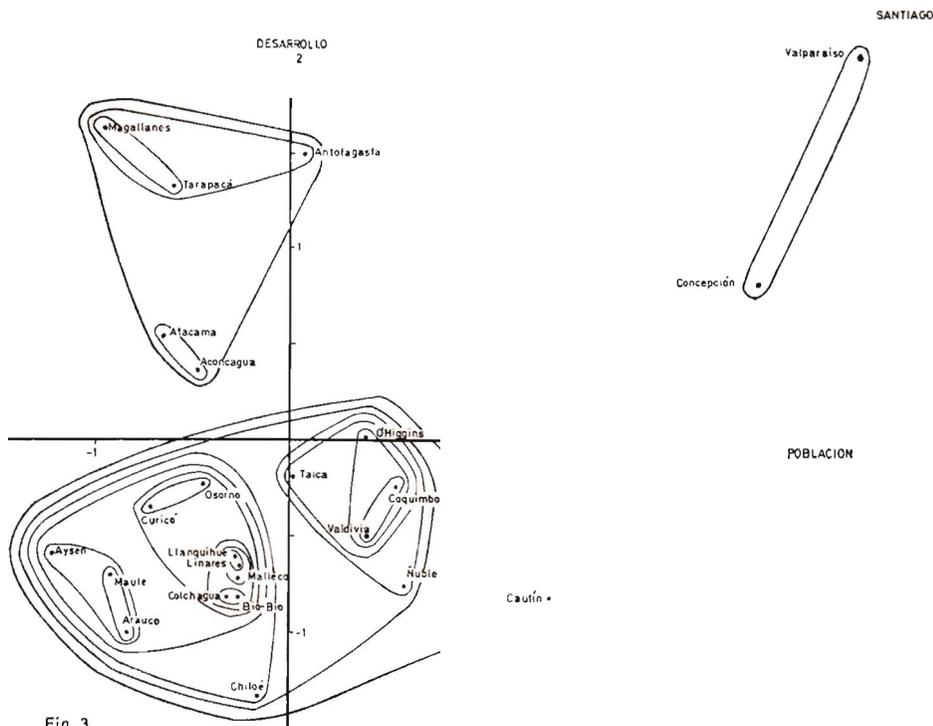


Fig. 3

los grupos contiene las diecisiete provincias rurales del centro y sur de Chile que son relativamente bajas en desarrollo.

Este grupo más grande puede volver a subdividirse en provincias menos desarrolladas con poblaciones bajas (Colchagua, Curicó, Maule, Linares, Arauco, Bio-Bío, Malleco, Osorno, Llanquihue, Chiloé y Aisen), de las cuales Chiloé es la menos desarrollada, y las provincias que tienen similares niveles de desarrollo, pero cuya población está sobre la media. El segundo subgrupo incluye Coquimbo, O'Higgins, Talca, Ñuble, Cautín y Valdivia, de los cuales Ñuble y Cautín tienen cifras más bajas de desarrollo. Estos dos subgrupos se unen en el paso 18, pero no es hasta que se llega al paso 21 que Cautín se incorpora. Cautín, con su alta población, es por lo tanto el miembro menos común del grupo menos desarrollado.

Después del paso 21, las cuatro distintas categorías de provincia se unen. Al grupo de provincias rurales del centro y sur se une primero el grupo del norte y extremo sur, y luego Valparaíso - Concepción. La provincia de Santiago permanece aparte del resto del país hasta el último paso.

EXTENSIONES DEL METODO

No se pretende que los cuatro grupos principales de provincias que se distinguen en esta clasificación sean los "correctos" que deberían usarse al planificar el desarrollo futuro del país. Pero cuando se esté considerando una futura asignación de recursos, una clasificación de las provincias según algún criterio específico, es una ayuda muy útil. El método de clasificación que se ha mostrado aquí es sencillo, pero eficiente y objetivo, y fácilmente puede usarse con cualquiera variable o variables.

Fue usado por primera vez en Geografía por Berry (2). Cuando es necesario clasificar provincias según más de dos variables, ya no es posible medir la distancia entre dos puntos con una regla en un gráfico de dos dimensiones, pero la distancia se puede calcular usando el teorema de Pitágoras. Primero se encuentra la diferencia entre cada par de medidas para cada variable que se está considerando. Entonces estas diferencias se elevan al cuadrado y se suman. La raíz cuadrada del resultado es la distancia que se busca. Este tipo de cálculo lo hace muy rápidamente una computadora.

Un problema geográfico común es el de regionalización, que implica la aglomeración de áreas vecinas que tienen características similares. Las regiones pueden definirse usando este mismo método de clasificación, sencillamente siguiendo la regla que un área puede unirse a un grupo sólo si es contigua con al menos algún otro miembro del grupo. Esto asegura que todos los grupos estén constituidos por áreas vecinas. Si se satisface esta limitación, las regiones resultantes serán aquellas que tienen las máximas semejanzas dentro de sus límites y la máxima diferencia entre ellas. Hay algunas situaciones en que es necesario trabajar con regiones contiguas (tal como para fines administrativos), pero para la investigación geográfica común el agrupamiento que se hace sin considerar si los grupos finales forman regiones contiguas o no es generalmente más eficiente e informativo. Si existe un agrupamiento regional, se puede confiar que el método de clasificación sin limitaciones lo revelará.

BIBLIOGRAFIA

BERRY, B. J. L.: "A method for Deriving Multifactor Uniform Regions". *Przegląd Geograficzny*, XXXIII, 1961, 263-82.

MATTELART, ARMAND y GARRETON, MANUEL: "*Integración Nacional y Marginalidad*". Santiago, Chile. Edit. del Pacífico, 1965. pp. 38-39.

(2) Berry, B. J. L. "A method for deriving multifactor uniform regions", *Przegląd Geograficzny*, XXXIII (1961) 263-82.