
PROGRAMA RECURSIVO EN PASCAL PARA DETERMINACION DE ZONAS EN UNA SECUENCIA, POR EL METODO DE ANALISIS DE VARIANZA

LUIS A. MONTES

Profesor Asistente

Depto de Geociencias-Facultad de Ciencias-Universidad Nacional de Colombia

Montes L.A.: Programa Recursivo en Pascal para determinación de zonas en una secuencia, por el método de Análisis de Varianza. Geofís. Colomb. 1:31-36, 1992. ISSN 0121-2974

RESUMEN

Los datos geológicos comúnmente forman una secuencia, como registros eléctricos de pozo, pudiendo dividirse en segmentos homogéneos. Las zonas pueden identificarse por métodos de análisis de varianza. El programa contiene un procedimiento recursivo que evalúa un coeficiente de zonación entre varios posibles arreglos de datos, seleccionando el mayor valor para seleccionar el límite entre los dos nuevos segmentos. Este proceso se repite en cada zona hasta que el tamaño del segmento alcance un límite.

En un registro eléctrico se identificaron 21 contactos. El programa puede ser usado en otras áreas de la Geología.

ABSTRACT

Geologic data commonly form a sequence, for example well logs, and they may be subdivided into relative uniform intervals corresponding to stratigraphic units. Distinct zones within ordered data sets can be identified by analysis of variance method. The computer program is a recursive procedure in which arrangements into two contiguous zones are considered at each stage. In each proposed boundary a zonation index is calculated, and is selected the highest one to establish two different zones. The procedure is repeated for each zone until number of data in segments reach a limit.

In a digitized electric log, it was possible to identify 21 boundaries.

The program is general and can be used in others branch of geology.

1. INTRODUCCION

Muchos datos geológicos se presentan como secuencias de datos. En ellas es posible reconocer intervalos con características comunes o zonas de litología constante.

Este proceso de dividir una secuencia en segmentos relativamente uniformes se denomina Zonación.

Las distintas zonas dentro de la secuencia pueden identificarse por el método estadístico

de Análisis de varianza.

La búsqueda de estos "contactos" implica una técnica recursiva, en la cual en la secuencia se evalúa un índice, que permite determinar dos segmentos, en los cuales nuevamente se realiza el proceso de determinación del índice y la posterior segmentación.

El proceso termina cuando los segmentos alcanzan un límite.

El método se implanta a través de un programa escrito en lenguaje Pascal, y se ensaya en una secuencia de un registro eléctrico de un pozo en la localidad de Madrid, C/marca.

2. METODO DE ZONACION

El método de Análisis de Varianza examina las diferencias entre las medias de dos poblaciones cuyos elementos contribuyen a través de un modelo lineal.

El modelo plantea que a toda señal contribuyen tres factores:

$$X(ij) = \mu + \sigma(i) + \epsilon(ij)$$

Donde

μ : Media de toda la población.

$\sigma(i)$: es el efecto de la zona particular.

$\epsilon(ij)$: error aleatorio asociado con cada observación.

El método se basa en dos mediciones diferentes e independientes de la Varianza.

La suma de cuadrados dentro de la secuencia total se calcula de acuerdo a la siguiente formula:

$$SSW = \sum_{j=1}^m \Sigma (X_{ij} - \bar{X}_j)^2 / \sum_{j=1}^m n_j - m \quad [1]$$

X_{ij} = i-punto en j-segmento

\bar{X}_j = media de j-segmento

n_j = # de puntos en j-segmento

m = # de segmentos

La suma de cuadrados entre segmentos se calcula como:

$$SSB = \sum_{j=1}^m (\bar{X}_j - \bar{X})^2 / (m-1) \quad [2]$$

\bar{X} = media total de la secuencia

$X(ij)$ = i-observación en j-segmento.

$\bar{X}(j)$ = media del j-segmento.

$n(j)$ = numero de observaciones en j-segmento.

m = numero total de segmentos.

Las formulas representan una medida de la varianza alrededor de la media total

\bar{X} de la secuencia.

El coeficiente de zonacion se calcula como :

$$R = 1 - SSW/SSB \quad [3]$$

3. PROCEDIMIENTO

Al principio la secuencia total se divide en dos segmentos, una parte inicial corta que constituye el primer segmento y la parte restante el segundo segmento.

A continuación se calculan SSW, SSB y el índice de zonacion R en cada segmento, según las formulas 1, 2 y 3.

El límite entre los dos segmentos se mueve a lo largo de la secuencia a posiciones sucesivas, de acuerdo a un valor de paso de avance definido,

por cada movimiento del límite se calculan nuevamente los valores de SSW, SSB y R.

Al terminar de barrer la secuencia se obtienen un conjunto de índices de Zonación; uno por cada posición de el límite.

Se busca dentro de ese conjunto el valor de R máximo, y se escoge dicha posición como la primera zona de contacto.

Luego cada nuevo segmento generado se particiona, considerándose cada segmento como un bloque completamente distinto, iterando el procedimiento, para conseguir nuevas zonas de contacto en cada segmento.

El proceso continua hasta que la secuencia se divide en segmentos que no pueden superar una longitud minima.

4. APLICACION

El programa se ensayó con un registro eléctrico de pozo digitalizado manualmente (Tabla 1); situado en la localidad de Madrid (Cundinamarca). Para este registro se utilizó una sonda corta normal de 20 cm y los valores se mostraron a profundidades mayores de 330 m. El programa se creó con el uso del Compilador HP Pascal-9000 y se ejecutó en un microcomputador Hewlett-Packard 9000, del Departamento de Geociencias de la Universidad Nacional.

Para la implantación de esta técnica se utilizó el lenguaje Pascal, por la facilidad que presenta este lenguaje de llamado recursivo de rutinas; una segunda razón es el hecho de ser Pascal un lenguaje muy conocido y de fácil aprendizaje. Su utilización permitió determinar 21 contactos en la secuencia del registro (ver Tabla 2.), los cuales coinciden intuitivamente con las inflexiones de la curva del registro (Fig.1 y 2).

Al observar la gráfica del registro de pozo, se puede observar la bondad de ajuste del método y del programa en la determinación de las zonas de contacto que encierran segmentos litológicamente homogéneos, siendo estos de diferentes espesores.

El comportamiento del programa está influenciado por el longitud del segmento inicial y por el valor del paso de avance al recorrer la secuencia, generando resultados diferentes para valores distintos de estos parámetros.

5. CONCLUSIONES

El uso de programas de computador en la determinación de estructuras geológicas, es cada día mas importante.

Se utilizó un programa para la identificación de zonas litológicamente homogéneas, y se pudo determinar su utilidad al aplicar el método en el análisis de un registro eléctrico.

Se pueden procesar secuencias de información extensas en periodos de tiempo muy cortos, el tiempo para analizar 100 datos le tomo a la maquina aproximadamente 20 segundos, incluido la generación de la salida de resultados. La técnica de zonacion se utiliza en otras áreas de la geología, que contemplen analisis de secuencias de datos.

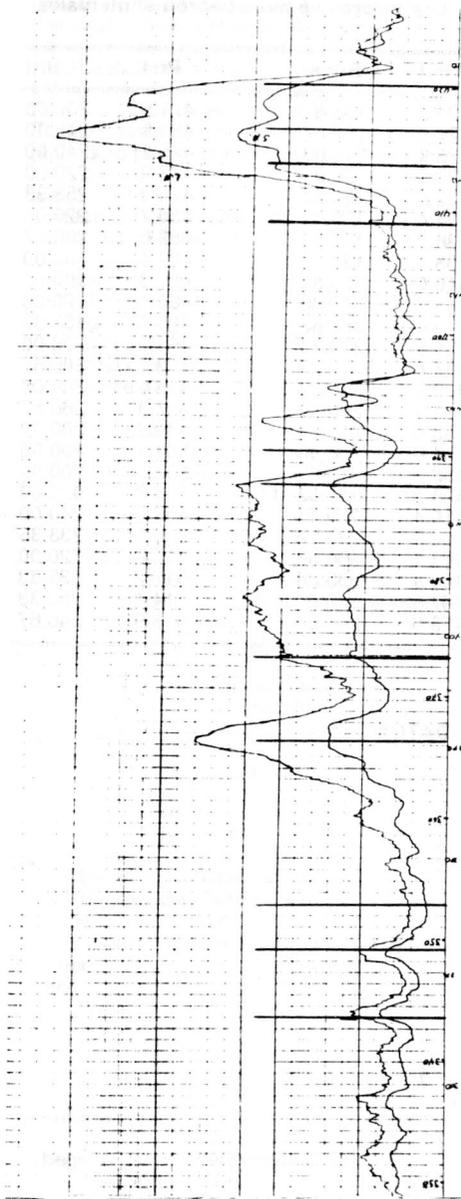


Figura 1. Registro eléctrico de pozo. Sensibilidad: 10 divisiones = 500 Ω m; Profundidad: cada división = 0.9434 m. Los segmentos indican unidades litológicas homogéneas. El tope de la columna es 332.6 m y la base 423.3 m.

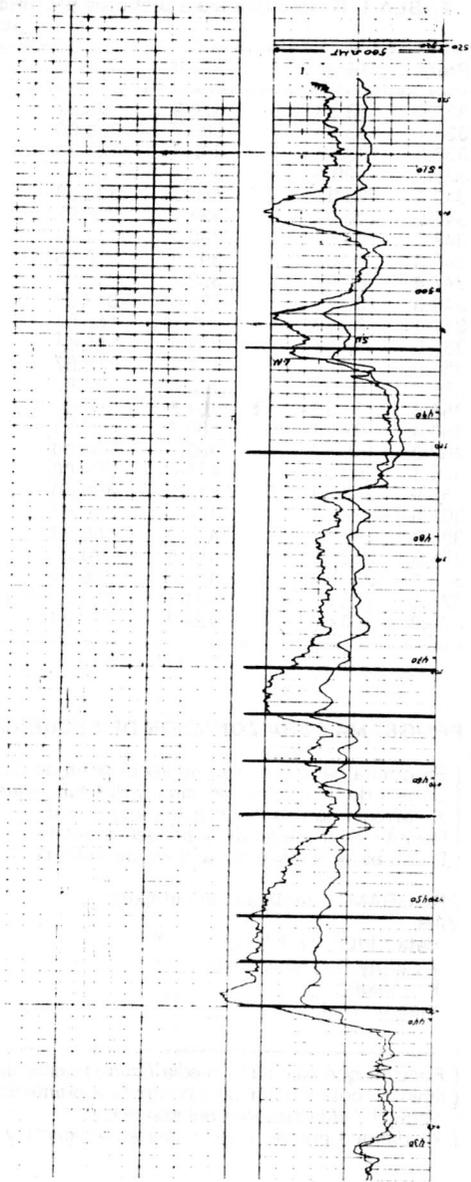


Figura 2. (continuación Fig.1). Registro eléctrico de pozo. Sensibilidad: 10 divisiones = 500 Ω m; Profundidad: cada división = 0.9434 m. Los segmentos indican unidades litológicas homogéneas. El tope de la columna es 425 m y la base 519 m.

TABLA 1. Resistividades a distintas profundidades. Los valores se muestrearon a intervalos iguales.

Prof	Resist	Prof	Resist	Prof	Resist	Prof	Resist
332.6	133.33	378.9	293.33	425.2	186.67	471.5	200.00
334.4	133.33	380.7	220.00	427.0	186.67	473.3	240.00
336.3	166.67	382.6	133.33	428.9	213.33	475.2	240.00
338.1	100.0	384.4	306.67	430.7	126.67	477.0	220.00
340.0	133.33	386.3	333.33	432.6	126.67	478.9	253.33
341.9	100.00	388.1	266.67	434.4	126.67	480.7	220.00
343.7	186.67	390.0	173.33	436.3	126.67	482.6	266.67
345.6	86.67	391.9	300.00	438.1	133.33	484.4	160.00
347.4	86.67	393.7	313.33	440.0	333.33	486.3	120.00
349.3	160.00	395.6	313.33	441.9	353.33	488.1	120.00
351.1	80.00	397.4	186.67	443.7	353.33	490.0	133.33
353.0	66.67	399.3	166.67	445.6	333.33	491.9	200.00
354.8	66.67	401.1	166.67	447.4	333.33	493.7	306.67
356.7	106.67	403.0	146.67	449.3	313.33	495.6	300.00
358.5	133.33	404.8	166.67	451.1	280.00	497.4	266.67
360.4	133.33	406.7	193.33	453.0	266.67	499.3	200.00
362.2	166.67	408.5	200.00	454.8	200.00	501.1	200.00
364.1	253.33	410.4	200.00	456.7	226.67	503.0	200.00
365.9	353.33	412.2	400.00	458.5	333.33	504.8	320.00
367.8	333.33	414.1	546.67	460.4	266.67	506.7	240.00
369.6	186.67	415.9	500.00	462.2	233.33	508.5	233.33
371.5	213.33	417.8	573.33	464.1	366.67	510.4	220.00
373.3	286.67	419.6	466.67	465.9	333.33	512.2	233.33
375.2	286.67	421.5	253.33	467.8	300.00	514.1	233.33
377.0	300.00	423.3	233.33	469.6	266.67	515.9	246.67

PROGRAMA PARA ZONACION DE SECUENCIAS DE DATOS

```
{ Programa para zonacion de secuencias de datos }
{ Consta de un modulo principal y dos funciones }
{ X : arreglo que contiene datos a procesar }
{ Meant : valor medio de la secuencia total }
{ Datos es el nombre del archivo de entrada }
```

```
PROGRAM zone(data,input, output);
```

```
VAR
```

```
data : FILE OF REAL;
x : array [1..256] of real;
k : integer;
meant : real;
```

```
{-----}
{ Función que calcula la media del segmento definido }
{ liminf1 : cota inferior del segmento a promediar }
{ limsup1 : cota superior del segmento }
{ X : datos a procesar, definidos entre liminf1 y limsup1 }
```

```
function mean(liminf1, limsup1 : integer) : real;
```

```
var
```

```
i : integer;
suma : real;
begin
suma := 0.0;
for i := liminf1 to limsup1 do
suma := suma + x[i];
mean:=suma/(limsup1 - liminf1);
end;
```

```
{=====}
```

```

{ Mean1 : media del primer segmento          }
{ Mean2 : media del segundo segmento         }
{ liminf, limsup : limites del segmento a analizar y dividir }
{ R : coeficiente a evaluar, para determinacion de limites }
{ Pos : indicador de R evaluado              }
{ Inc : longitud de paso de avance           }
{ Delta : indicador de posición de máximo R }
}
procedure border(liminf, limsup : integer);

```

```

var
  mean1, mean2, ssw, ssb, rmax, r : real;
  pos, inc, i, delta : integer;

begin
  if (limsup - liminf) > 8 then
    begin
      pos := 0; inc := 4;
      rmax := 0.0;
      while (limsup-liminf-inc > 4) do
        begin
          ssw := 0;
          mean1 := mean(liminf, liminf+inc);
          mean2 := mean(liminf+inc+1, limsup);
          for i := liminf to liminf+inc do
            ssw := ssw + sqr(x[i]-mean1);
          for i := liminf+inc+1 to limsup do
            ssw := ssw + sqr(x[i]-mean2);
          ssw := ssw/(limsup-liminf-4);
          ssb := sqr(mean1-mean2)+sqr(mean2-mean1);
          pos := pos + 1;
          if ssb > 0 then
            begin
              r := 1-(ssw/ssb);
              if r > rmax then
                begin
                  rmax:=r;
                  delta:=pos;
                end;
            end;
          inc := inc + 2;
        end;
      border(liminf+delta+1, limsup);
      border(liminf, liminf+delta);
      if rmax > 0 then
        writeln(liminf+delta, ' ',rmax:4:3);
    end;
  end;

```

```

(=====)

```

```

begin { ** inicio del principal ** }
  reset(data,'datos');
  k := 0;
  while (not Eof(data)) do
    begin
      k := k + 1;
      read(data, x[k]);
    end;
  close(data);
  meant:=mean(1,k);
  border(1,k);
end. { ** fin del principal ** }

```

Tabla 2. Muestra el tope superior del segmento y el Coeficiente R

Cota superior del segmento	R	Profundidad (m)
88	0.323	493.6
84	0.223	486.3
74	0.573	467.8
72	0.785	465.9
70	0.807	460.4
68	0.462	456.7
63	0.802	447.4
61	0.879	443.7
59	0.423	440.0
54	0.081	430.7
48	0.790	419.6
46	0.940	415.9
44	0.918	412.2
42	0.361	408.5
32	0.155	390.0
30	0.221	386.3
23	0.268	373.3
19	0.148	365.9
12	0.442	353.0
10	0.484	349.3
7	0.389	343.7

BIBLIOGRAFIA

- Davis, J. C.** (1986): Statistics and data analysis in Geology. 211pp. John Wiley & Sons Co, New York.
- Gill, D.** (1970): Application of statistical zonation method to reservoir evaluation and digitized-Log Analysis. The Am. Assoc. of Petrol. Geologists Bulletin. **54** (5):719-729.
- Miller, R. L. & J.S. Kahn.** (1962): Statistical analysis in the geological sciences. John Wiley & Sons Co.
- Tucker, A.B.** (1987): Programming languages. pp.18-66. McGraw-Hill Inc, New York.