

APLICAÇÃO DE FUNÇÕES DISCRIMINANTES NA DISTINÇÃO DE ASSINATURAS ESPECTRAIS
DE CÁLCARIOS E QUARTZITOS EM IMAGENS MSS-LANDSAT

J. T. de Mattos e I. Vitorello

Instituto de Pesquisas Espaciais

Conselho Nacional de Desenvolvimento Científico e Tecnológico

Caixa Postal, 515 - 12200 - São José dos Campos, SP, Brasil

RESUMO

Este trabalho objetiva aplicar técnicas de estatística multivariante utilizando-se funções discriminantes sobre dados de fitas compatíveis com Computador (CCTs) do sistema LANDSAT, na análise de separabilidade entre calcários e quartzitos da região da Serra da Canastra, no estado de Minas Gerais (MG), bem como no estabelecimento da contribuição de cada canal do MSS-LANDSAT na separação das duas litologias citadas. Como áreas de estudo selecionaram-se diferentes porções das imagens correspondentes a órbita 164, ponto 26 de 21/06/76. Extraiu-se a resposta espectral correspondente a várias áreas de treinamento de uma mesma litologia, anotando-se os parâmetros estatísticos fornecidos pelo programa de computador "Single Cell", que utiliza o algoritmo de classificação do tipo determinístico, conhecido como Método do Paralelepípedo. Utilizou-se também o programa "Slicer", divisão de canal em intervalos de intensidade, para estabelecer e avaliar a expressão do comportamento espectral dos alvos analisados. Para a aplicação da técnica de funções discriminantes selecionaram-se como "amostragem espectral" os dados de variâncias das duas litologias estudadas, por representarem medidas de dispersão, facilitando assim a aplicação do método de regressão linear. Como teste de significância da função discriminante empregou-se o método da Distância Generalizada de Mahalanobis, o que permitiu caracterizar a discriminação entre os quartzitos e calcários analisados.

ABSTRACT

The objective of this study was to apply Multivariate Statistical Techniques, using discriminant functions on CCTs (Computer Compatible Tapes) from the LANDSAT system, to separate limestones and quartzites in the area of Serra da Canastra, in the state of Minas Gerais. Furthermore, the intention was also to establish the contribution of each MSS-LANDSAT spectral band to separate these geological units. Spectral signatures and respective statistics of both geological units were obtained from "training areas" within selected "study areas", from the four spectral bands of a 06.26.76 LANDSAT passage, (orbit 164, point 26), and through the use of a Parallelepiped Classifier (Single-Cell Signature Acquisition). The technique of Density Slicing was then utilized in the evaluation of the spectral signatures of the targets. The population variances of the spectral distribution of both lithologies were chosen as spectral samples for the discriminant function, since they represent a dispersion measure, and thus ease the application of the linear regression method. The Mahalanobis Distance method was used as the test of significance for the Discriminant Function, allowing the characterization of the separability between quartzites and limestones.

1. INTRODUÇÃO

O emprego das técnicas estatísticas multivariantes, utilizando-se de funções discriminantes, tem sido muito aplicado à análise espectral de dados de sensoriamento remoto, principalmente aqueles que permitem digitalização, constituindo-se numa ferramenta valiosa para a redução de grande volume de dados e facilitando sua interpretação.

O objetivo deste trabalho é mostrar como observações de dados multivariados podem ser reduzidos a poucos membros finais hipotéticos, de tal forma que as amostras sejam agrupadas ou classificadas de acordo com a proporção de cada um de seus membros. Para tanto, foram analisadas imagens MSS-LANDSAT sob forma de fitas CCTs e utilizado o analisador multiespectral IMAGE-100 (General Electric Company, 1975) sist

tema computacional de classificação multiespectral de assinaturas de alvos naturais.

Selecionou-se como área de estudo a região da Serra da Canastra e Represa de Furnas, situada no sul do Estado de Minas Gerais, visto que esta região apresenta grandes exposições de litologias distintas e os autores têm vários conhecimentos de campo sobre a área. Em função disso, são aqui constatadas e analisadas as variações no registro espectral de duas unidades litológicas, significativas da área, (calcários e quartzitos) cuja intenção é, ao confrontar o comportamento espectral de distintos alvos entre si, mostrar a discriminação que ocorre quando se aplicam técnicas de análise discriminante.

2. AMOSTRAGEM DOS DADOS

Os dados empregados correspondem a uma amostragem do comportamento espectral de diferentes áreas selecionadas da imagem MSS do LANDSAT de nº 176178, órbita 164, ponto 26, com tamanhos uniformes de 512 por 512 "pixels". Na escolha das áreas com estas dimensões evita-se modificar os dados originais do satélite e com isto é possível obter o máximo de informação radiométrica contida na fita, ficando também as áreas geométrica e estatisticamente corretas.

Em cada área de estudo, utilizando-se os quatro canais do MSS (dois na região do visível e dois na região de infravermelho próximo), extraem-se as respostas espectrais correspondentes às várias áreas de treinamento de uma mesma unidade geológica, anotando-se seus parâmetros estatísticos (técnica de classificação do tipo "Single-Cell Signature").

Nesta fase de amostragem, a aquisição das assinaturas espectrais, através dos histogramas de frequência de "pixels" de cada nível do intervalo de resposta do alvo estudado, em todos os canais, também foi obtida através da técnica de "slicer" (General Electric Company, 1974) que permite dividir um canal MSS escolhido em intervalos de níveis de cinza que mais caracterizem um determinado alvo.

A obtenção de respostas espectrais (intervalo de variação tonal, variância e limites de "range" espectral), nesta fase de amostragem, teve como objetivo caracterizar os diversos alvos geológicos (unidades litológicas), visando estabelecer e avaliar a expressão (extensão) de suas assinaturas para, assim, aplicar as técnicas de análise discriminante somente nos alvos ou unidades geológicas com comportamento expressivos em área, áreas estas com diferentes distribuições geográficas.

3. MÉTODO E PROCEDIMENTO ANALÍTICO

A função discriminante é uma técnica multivariante utilizada quando se deseja transformar um conjunto de medidas num único valor R_p , definido como uma combinação linear dos valores originais.

O desenvolvimento desta técnica (Davis, 1973) parte do método de regressão linear, empregando-se para a solução de sistemas de equações lineares, o cálculo matricial segundo:

$$[V_p^2] [\lambda_p] = [R_p]$$

onde:

$[V_p^2]$ = matriz $p \times p$ das variâncias e covariâncias combinadas das "p" variáveis;

$[\lambda_p]$ = coeficientes da função discriminante;

$[R_p]$ = diferença entre as médias das variáveis de "p" grupos.

No cálculo de $[V_p^2]$ determinam-se a matriz da soma de quadrados e produtos cruzados de todas as variáveis dos grupos analisados (neste caso 2 grupos: Quartzitos e Calcários) e as médias de cada coluna dos grupos, a fim de obter os $[R_p]$. Para a obtenção dos coeficientes $[\lambda_p]$ emprega-se a equação: $[\lambda_p] = [V_p^2]^{-1} [R_p]$. Finalmente calculam-se os valores centrais do 1º grupo (D_A) e do 2º grupo (D_B) e o índice de discriminante (D_0).

Como teste de significância da função encontrada, empregou-se o Método da Distância Generalizada de Mahalanobis (D^2), utilizando-se esta medida para ser testada pela distribuição F.

4. RESULTADOS

No trabalho foram utilizadas duas unidades litológicas, quartzitos e calcários, para a aplicação do método de funções discriminantes, visto que estas litologias apresentam expressão de resposta espectral global, servindo portanto para classificação automática em grandes áreas.

Empregaram-se como dados de amostragem os valores de variância da resposta espectral, do dois alvos, para cada canal, pois estes valores representam medidas de separabilidade (desvio) entre amostras. Nas Tabelas 1 e 2 acham-se listados os dados obtidos.

A obtenção das respostas espectrais permitiu a limitação do número de amostras analisadas, em virtude de se obter, nas áreas testadas, resposta espectral significativamente percentente a um único alvo. Desse modo, de todas

as "áreas testes", somente 15 foram selecionadas para a análise (sendo 5 para as regiões calcárias e 10 para as de quartzitos), as quais

apresentam "ranges" espectrais bem definidos e estão relacionadas com os principais pontos de ocorrência das unidades analisadas.

TABELA 1

DADOS DE VARIÂNCIA DA RESPOSTA ESPECTRAL DE ROCHAS CALCÁRIAS DOS 4 CANAIS DO LANDSAT, OBTIDOS NO IMAGE-100

VARIÂNCIA (CALCÁRIOS) VARIÁVEIS-C (CANALIS-LANDSAT)	Ca ₁	Ca ₂	Ca ₃	Ca ₄	Ca ₅
C ₁ (4)	2,6	3,0	4,1	3,0	6,0
C ₂ (5)	7,9	9,5	10,6	7,7	8,1
C ₃ (6)	20,4	19,4	26,6	13,0	32,7
C ₄ (7)	68,6	58,5	68,5	48,2	76,7

TABELA 2

DADOS DE VARIÂNCIA DA RESPOSTA ESPECTRAL DE ROCHAS QUARTZÍTICAS DOS 4 CANAIS DO LANDSAT, OBTIDOS NO IMAGE-100

VARIÂNCIA (QUARTZITOS) VARIÁVEIS-C (CANALIS-LANDSAT)	Q ₁	Q ₂	Q ₃	Q ₄	Q ₅	Q ₆	Q ₇	Q ₈	Q ₉	Q ₁₀
C ₁ (4)	2,88	7,88	9,71	4,73	4,00	3,90	2,30	4,90	4,20	5,10
C ₂ (5)	6,27	16,64	18,66	11,33	10,00	8,20	4,90	12,40	13,10	16,60
C ₃ (6)	8,57	23,38	24,09	22,45	29,00	19,40	25,50	18,60	43,80	48,10
C ₄ (7)	19,10	42,6	36,40	35,92	37,40	15,60	42,00	28,50	53,00	58,20

Desta forma, pôde-se realizar a análise discriminante, procurando-se correlacionar a separabilidade obtida para as litologias existentes, comparando-se os resultados com os diversos canais MSS analisados. Os dados que se seguem correspondem aos cálculos realizados segundo a análise discriminante (Tabela 3) e podem ser obtidos através de qualquer sistema computacional.

Teses das Hipóteses

$$H_0 \rightarrow \mu_A - \mu_B$$

As duas médias multivariante são iguais (as distância entre os grupos são iguais).

Se q (função discriminante) for menor que F_α (função de distribuição - tabelado) H_0 será aceito.

$$F_\alpha = (4; 10 + 5 - 4 - 1) \rightarrow F_\alpha = (4; 10)$$

Para $\alpha = 0,05$

$$F_{0,05; (4; 10)} = 3,48$$

e

$$q = F \text{ (calculado)} = 15,83$$

TABELA 3

SEQUÊNCIA DE DADOS PARA EXEMPLO DA APLICAÇÃO DE ANÁLISE DISCRIMINANTE

<u>NÚMERO DE OBSERVAÇÕES = 10 (QUARTZITOS - GRUPO 1)</u>						
MATRIZ DA SOMA DOS QUADRADOS DO GRUPO 1						
4.518980000	01					
8.187430000	01	1.901860000	02			
1.580400000	01	2.188341000	02	1.238904690	03	
5.130040000	01	2.817754000	02	1.258834920	03	1.61394250 03
MÉDIAS DE CADA COLUNA DO GRUPO 1						
4.960000000	00	1.181000000	01	2.628900000	01	3.68720000 01
<u>NÚMERO DE OBSERVAÇÕES = 5 (CALCÁRIOS - GRUPO 2)</u>						
MATRIZ DA SOMA DOS QUADRADOS DO GRUPO 2						
7.632000000	00					
3.880000000	-01	6232000000	00			
3.624600000	01	1039400000	01	2.250880000	02	
3.724000000	01	-9780000000	00	2.637200000	02	4.74540000 02
MÉDIAS DE CADA COLUNA DO GRUPO 2						
3.740000000	00	8.7600000000	00	2.242000000	01	6.21000000 01
<u>MATRIZ DE COVARIÂNCIA</u>						
4.063215385	00					
6.327869231	00	1.510907692	01			
4.003846154	00	1.763293077	01	1.126148223	02	
6.810800000	00	1.092272308	01	1.171196092	02	1.606525046 02
<u>COEFICIENTES (λ_p)</u>			<u>VALORES</u>			
8.781907109	- 01 (canal 4)		DA =	3.5767	VALORES CENTRAIS	
-3.791412658	- 02 (canal 5)		DB =	28.2701	ÍNDICE DISCRIMINANTE	
8.517754027	- 01 (canal 6)		DO =	-15.9234	D.MAHALNOBIS	
-8.102924828	- 01 (canal 7)		D ² =	24.6933	DISTRIBUIÇÃO F	
			F =	15.8291		
<u>VALORES D(i) (PONTOS PARA DESCRIMINAÇÃO DOS GRUPOS)</u>						
GRUPO (A) QUARTZITOS			GRUPO (B) CALCÁRIOS			
- 5.8854	Q ₁		- 36.2261	Ca ₁		
- 8.3147	Q ₂		- 28.6033	Ca ₂		
- 1.1556	Q ₃		- 21.5462	Ca ₃		
- 6.2591	Q ₄		- 25.6404	Ca ₄		
- 2.4698	Q ₅		- 29.3343	Ca ₅		
6.9979	Q ₆					
-10.4780	Q ₇					
- 3.4173	Q ₈					
- 2.4460	Q ₉					
- 2.3392	Q ₁₀					

Do teste realizado, como F calculado (15,82) é maior que o tabelado $F_{\alpha}(3,48)$ a hipótese H_0 é rejeitada, concluindo-se que as médias das variâncias populacionais de A (quartzitos) são diferentes das médias das variân-

cias populacionais de B (calcários), portanto existe discriminação entre as populações.

A Figura 1 representa a projeção gráfica da função discriminante.

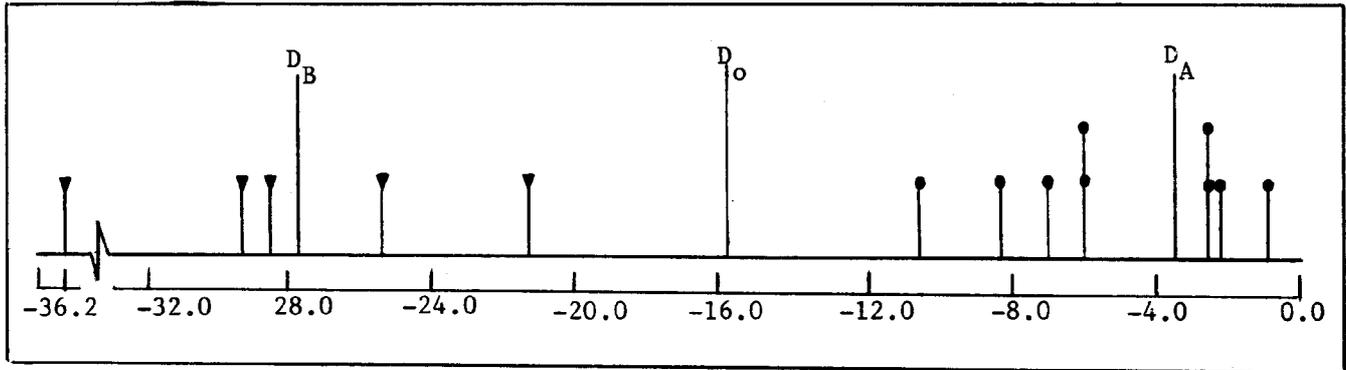


Fig. 1 - Representação gráfica da função discriminante : (D_A) - projeção dos valores centrais das médias de quartzitos; (D_B) - projeção dos valores centrais das médias dos calcários e (D_0) índice discriminante. A diferença entre D_A e D_B corresponde à distância Mahalanobis (D^2)

A contribuição relativa, em porcentagem, de cada variável (canal-LANDSAT) para o distanciamento entre os dois grupos foi obtida através

de $[C = \lambda_p (\bar{x}_A - \bar{x}_B) / D^2]$ cujos resultados são apresentados na Tabela 4.

TABELA 4

VALORES EM % DA SEPARABILIDADE DOS ALVOS NOS 4 CANAIS DO LANDSAT

VARIÁVEL (CANAL LANDSAT)	DIFERENÇA ENTRE MÉDIAS POPULACIONAIS $\bar{x}_A - \bar{x}_B$	COEFICIENTES (λ_p)	CONTRIBUIÇÃO %
1(4)	1.22	0.878	4,33
2(5)	3.05	-0.038	-0,47
3(6)	3.86	0.851	13.3
4(7)	-25.23	-0.816	82.7 *
$D^2 =$ Distâncias Mahalanobis = 24.693			
* = Maior separabilidade no canal 7 do MSS-LANDSAT			

5. CONCLUSÕES

Empregando-se a análise discriminante é possível estabelecer critérios de separabilidade de dados espectrais intercorrelacionados a alvos discretos, que se associam ao substrato geológico.

Apesar do pequeno número de amostras analisadas, o método revelou-se de grande utilidade na discriminação espectral quantitativa dos alvos e na constatação do recobrimento entre as amostras de áreas diferentes. Deve-se portanto ser testado em um maior número de

ocorrida em termos de confiabilidade (intervalo de confiança) bem como a separabilidade em cada canal do LANDSAT.

Como a utilização do método presuppõe que as variáveis (no caso, canais do LANDSAT) tenham distribuição normal, o uso de programas ("Clustering", Média K, Seleção de atributos, etc) que forneça valores normalizados são os mais aconselháveis para a aplicação da análise discriminante e, principalmente, para a obtenção dos parâmetros estatísticos, como a variân-

No caso específico dos quartzitos e calcários analisados verificou-se que estas duas litologias mostram grande separabilidade entre suas respostas espectrais e que o canal 7 (região do infravermelho próximo) é o que dá maior contribuição na separabilidade entre elas. Também pode-se observar, através da visualização gráfica (Figura 1) que as amostras de quartzitos acham-se mais agrupadas que as de calcário, possuindo portanto "ranges" espectrais mais estreitos, o que lhe confere assinaturas espectrais mais distintas e do tipo extensão global (alvos com assinaturas bastante significativas em escala regional).

Além disso, a perfeita concordância entre a individualização destas unidades por métodos convencionais com a fornecida pelo método estatístico reforça a utilização da técnica na quantificação dos dados espectrais para fins geológicos.

6. REFERÊNCIAS BIBLIOGRÁFICAS

DAVIS, J.C. Statistics and Data Analysis in Geology. *John Weley & Sons, New York: 442-456, 1973.*

GENERAL ELECTRIC COMPANY. Image 100 System Software Description. *Florida, Space Division, Ground System Department, 1974.*

— Image 100 System Description - Interactive Multispectral Image Analysis Systems. *Florida, Space Division, Ground Systems Department, 1975.*

7. BIBLIOGRAFIA COMPLEMENTAR

GOODENOUGH, SHLIENS, S. Automatic Classification Methodology. *Ottawa Canadá Center for Remote Sensing, Department of Energy Mines and Resources, 1974.*

KRUMBEIN, W.C.; GRAYBILL, F.A. An introduction to statistical models in geology. *Mc Graw Hill, New York: 475 p. 1965*

WATSON, R.D. Spectral Reflectance and Photometric Properties of Selected Rocks. *Remote Sensing of Environment, 2(2):95-100, 1972.*