

Tema 9

Clasificación

En el tema anterior se ha visto como pueden obtenerse medidas cuantitativas de una serie de factores que condicionan la respuesta espectral de los píxeles. Este enfoque responde a la idea de que existe un continuo y una transición suave entre los diferentes valores de reflectividad. La clasificación, por el contrario, responde a la idea de que pueden establecerse límites precisos entre las respuestas espectrales de unos píxeles y otros, siendo estos límites los que definen las clases.

Un ejemplo sencillo sería pensar en un experimento consistente en medir el peso y la altura del conjunto de alumnos de la Universidad de Murcia y representar los valores medidos en un gráfico con el peso en abscisas y la altura en ordenadas. ¿Aparecerían todos los puntos dispersos de manera que hubiese una transición suave entre las diferentes complejiones, o bien los puntos se aglomerarían en áreas del gráfico definiendo clases (elfos, orcos, enanos, hobbits y hombres)? En definitiva el problema es similar a la dicotomía entre objetos y superficies que aparece en el trabajo con SIG.

9.1 Introducción a las técnicas de clasificación. Generación de clases

La elaboración de una clasificación del área de estudio en una serie de clases relativas a litología, tipos de vegetación, usos del suelo, etc., es uno de los objetivos fundamentales en teledetección.

La clasificación en teledetección es, por tanto, un caso particular del problema general de clasificar N individuos procedentes de una muestra en un conjunto de $M < N$ clases en función de una serie de K variables (X_1, X_2, \dots, X_K) . Este problema se resuelve en dos fases:

1. Determinación del número de clases y de las propiedades de estas en relación a las K variables
2. Asignar cada uno de los N individuos a una de las M clases utilizando una regla de decisión basada en las propiedades de los individuos y las clases en relación a las K variables.

En teledetección el conjunto de variables está compuesto por la reflectividad en cada una de las bandas. Sin embargo, además de esta *información espectral* puede utilizarse *información textural e información contextual*. La información textural hace referencia a la características en la vecindad de un pixel. Se definen una serie de variables (reflectividad media, varianza, autocorrelación, etc.) que tratan de cuantificar algunas de las propiedades cualitativas que se estudian en fotointerpretación. El contexto por su parte hace referencia a la necesaria coherencia en cuanto a la clasificación de píxeles vecinos. No tendría, por ejemplo, mucho sentido clasificar un pixel como cultivo de invernadero al lado de píxeles clasificados como roquedo de alta montaña.

La incorporación de información textural y contextual tiene, sin embargo, el inconveniente de introducir variables que no suelen seguir una distribución gaussiana con lo que algunos métodos de clasificación no funcionan correctamente.

La determinación de las clases puede hacerse básicamente mediante dos criterios:

1. **Clasificación supervisada.** Se parte de un conjunto de clases conocido a priori. Estas clases deben caracterizarse en función del conjunto de variables mediante la medición de las mismas en individuos cuya pertenencia a una de las clases no presente dudas (*áreas de entrenamiento*). En la figura 9.1 se observa la caracterización espectral de un conjunto de clases de uso de suelo definidas a partir de los valores de reflectividad en las diferentes bandas de landsat MSS.
2. **Clasificación no supervisada.** No se establece ninguna clase a priori, aunque es necesario determinar el número de clases que queremos establecer, y dejar que las defina un procedimiento estadístico.

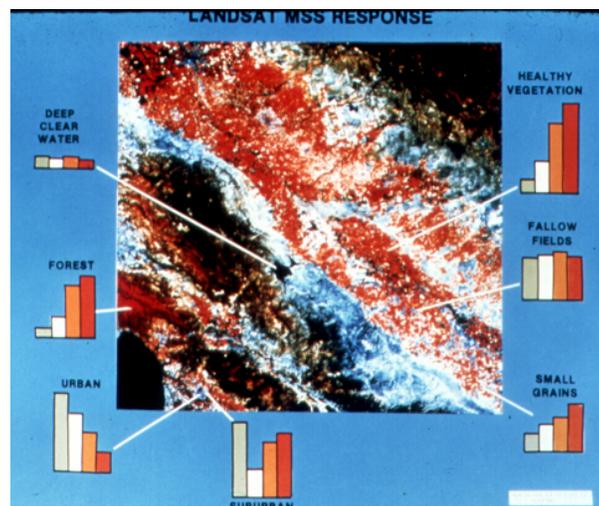


Figura 9.1: Imagen MSS y definición de clases como histogramas

La asignación de cada uno de los individuos a las clases obtenidas anteriormente puede hacerse mediante diferentes procedimientos:

- No estadísticos (árboles de decisión, mínima distancia, paralelepípedos)
- Estadísticos clásicos (máxima probabilidad)
- Algoritmos basados en inteligencia artificial (lógica borrosa, redes neuronales)
- Algoritmos que utilizan información contextual

9.2 Interpretación geométrica de la clasificación

Uno de los modos más sencillos de entender la clasificación cuando se basa en medidas cuantitativas es pensar en el conjunto de variables que caracterizan a individuos y clases como un conjunto de ejes que definen un *espacio de variables* multidimensional (figura 9.2). A partir de este momento, por simplificar, se presentará un ejemplo en el que se trabaja con las bandas 3 y 4 de landsat (rojo e infrarrojo cercano) para clasificar un conjunto de píxeles (figura 9.3).

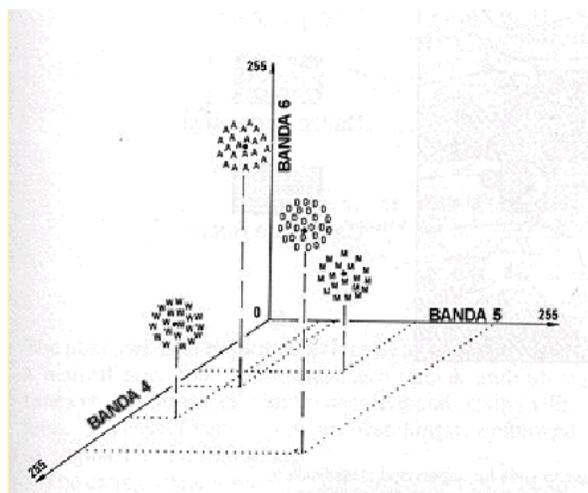


Figura 9.2: Definición de clases en el espacio de variables

La definición de clases puede hacerse de dos modos alternativos:

- Se divide el espacio de variables en regiones mediante el trazado de fronteras entre las mismas, cada región corresponde a una clase. Este procedimiento se utiliza en los árboles de decisión (figura 9.9), el método de los paralelepípedos y las redes neuronales.

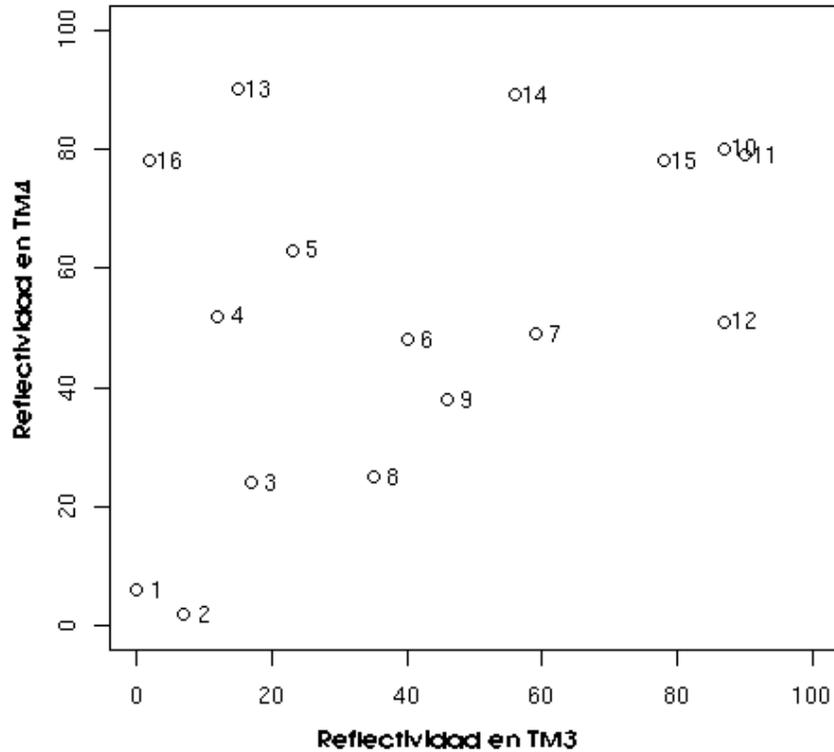


Figura 9.3: Datos de reflectividad en las bandas TM3 (x) y TM4 (y)

- Se asigna a cada clase un punto canónico que constituye el centroide de la clase. Cuanto más cerca este un individuo de este punto canónico más probable es que pertenezca a esta clase (figura ??). Se necesita por tanto una medida de distancia que equivale a una medida de semejanza o diferencia entre los diferentes individuos y entre los individuos y las clases. Dos individuos muy parecidos pertenecerán con toda seguridad a la misma clase, mientras que dos individuos distintos pertenecerán a diferentes clases. La medida más utilizada es la distancia euclidiana generalizada para espacios multidimensionales (aunque pueden utilizarse otro tipo de distancias):

$$d_{a,b} = \sqrt{\sum_{i=1}^N (X_{i,a} - X_{i,b})^2} \quad (9.1)$$

este enfoque es utilizado por el procedimiento de mínima distancia y el de máxima probabilidad.

agruparse sin dificultad.

En las figuras 9.4 y 9.5 puede verse con más detalle como funciona el clustering con los datos del ejemplo. Los individuos más próximos son el 1 y el 2 (distancia=8.06), por tanto se unen en la primera clase. A partir de ese momento, los individuos 1 y 2 ya no existen y se sustituyen por la clase 1-2 que queda definida por el círculo rojo intermedio. De este modo se van aglomerando individuos, individuos con clases (por ejemplo cuando se une la clase 6-9 con el individuo 7) o clases con clases (cuando se une la clase 1-2 con la 3-8). El procedimiento continua hasta la fusión de todas las clases en una sola o hasta alcanzar el número de clases establecido a priori. En la figura 9.5 aparece cual sería la clasificación resultante si se hubiesen escogido 4 clases.

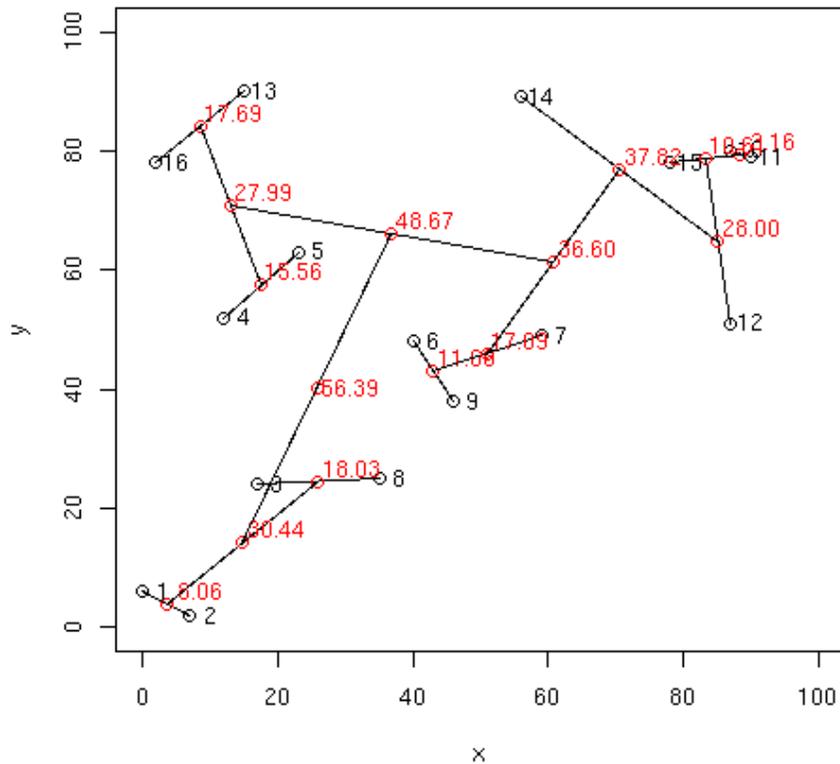


Figura 9.5: Análisis de cluster jerárquico

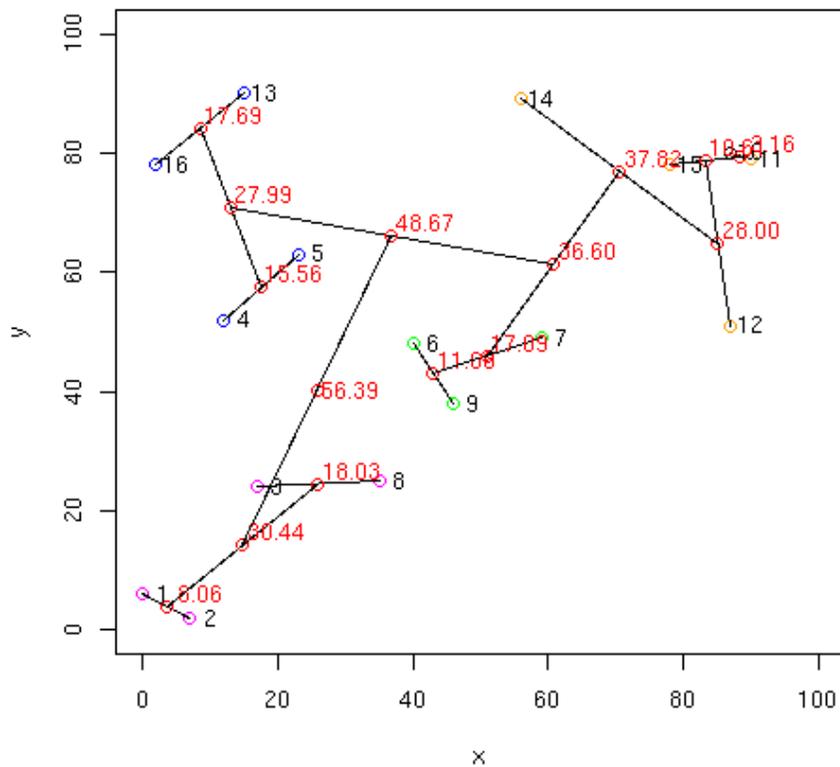


Figura 9.6: Clasificación no supervisada resultante

9.4 Clasificación supervisada

La clasificación supervisada se basa en la disponibilidad de *áreas de entrenamiento*. Se trata de áreas de las que se conoce *a priori* la clase a la que pertenecen y que servirán para generar una signatura espectral característica de cada una de las clases. Se denominan *clases informacionales* en contraposición a las *clases espectrales* que genera la clasificación no supervisada. Las áreas de entrenamiento deben ser lo más homogéneas posibles y en las que sepamos lo que había el día que se tomó la imagen. Por ello esta operación se realiza el mismo día en el que el satélite toma la imagen y luego se compra esta. Otra posibilidad es utilizar fotografía aérea o información de otro tipo.

Siguiendo con el ejemplo anterior, en la figura 9.7 los píxeles representados con tres colores diferentes corresponden a tres clases diferentes obtenidas a partir de las áreas de entrenamiento.

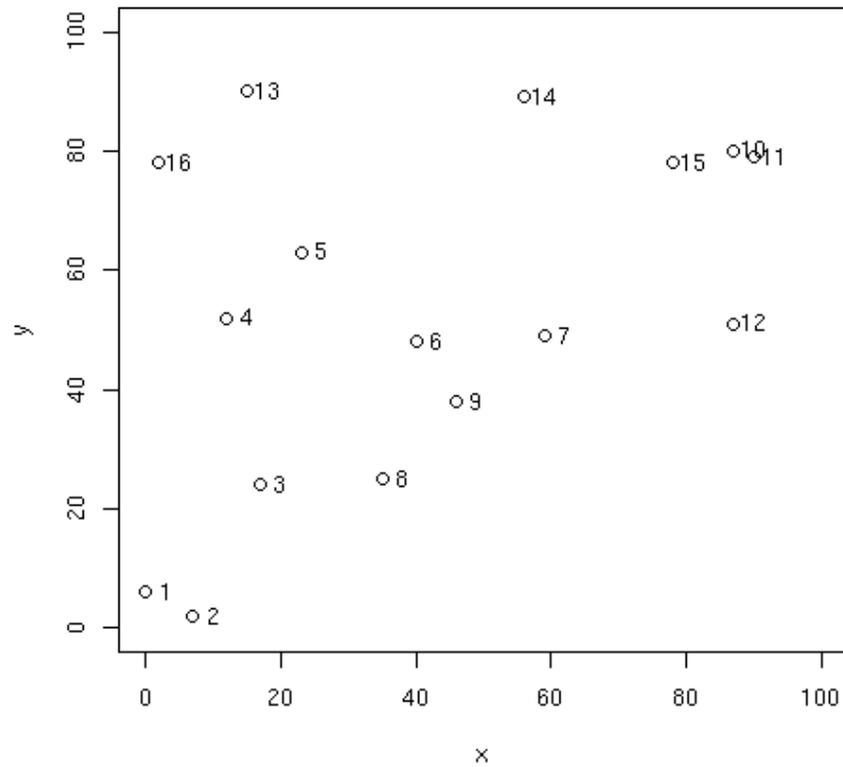


Figura 9.7: Incorporación de las areas de entrenamiento

Sin embargo, puesto que hemos creado las firmas espectrales con varios pixels de una misma clase, lo que obtenemos no es una firma en la que a cada banda se asigna un valor de reflectividad, sino una distribución de reflectividades para cada banda con una serie de estadísticos, los más relevantes para el análisis posterior son:

- media
- desviación típica
- máximo
- mínimo

9.5 Adjudicación de píxeles a clases. Técnicas no estadísticas

Una vez que se dispone de un conjunto de clases y de sus signaturas espectrales características, vamos a tratar de adjudicar cada uno de los píxeles a una clase. Los primeros métodos que se utilizaron eran de tipo no estadístico y con ellos se obtenían resultados bastante mediocres. Su análisis es, sin embargo, útil para comenzar a indagar en los problemas del proceso de clasificación.

En el caso de la clasificación no supervisada, hay que hacer la salvedad de que, en teoría no debiera ser necesaria esta segunda fase y a que en el proceso de clustering todos los individuos quedan asignados a una clase. Sin embargo el clustering en teledetección utiliza sólo una muestra de los píxeles, debido al gran número de píxeles existentes en una imagen, por tanto es necesario asignar los restantes de algún modo.

9.5.1 Árboles de clasificación

Es una técnica similar a la de las guías de identificación de rocas, árboles, etc. Se van planteando una serie de cuestiones respecto al píxel a clasificar y en función de las respuestas se hacen otras preguntas. Las cuestiones planteadas hacen referencia a los valores de reflectividad en las diferentes bandas de modo que, desde un punto de vista geométrico, trazan líneas que definen fronteras entre clases. Tras pasar a través de diferentes cuestiones respecto al píxel, este queda clasificado.

En el ejemplo que nos ocupa el árbol de decisión aparece en la figura ???. La división del espacio de variables a que da lugar este árbol de decisión aparece en la figura 9.9. Resulta evidente que la definición de las cuestiones que se plantean en un árbol de decisión es bastante subjetiva, aún teniendo en cuenta los valores obtenidos en las áreas de entrenamiento.

9.5.2 Clasificación por mínima distancia

Puesto que cada clase dispone de una reflectividad media para cada una de las bandas, podemos utilizar estas medias para situar cada una de las clases en el espacio de variables. Posteriormente, y en este espacio de variables se calcula la distancia entre un píxel cada una de las clases. Evidentemente cada píxel se asignará a la clase respecto a la cual su distancia sea mínima (figura ???).

Este método no es demasiado bueno ya que sobreclasifica la imagen, es decir ningún píxel queda sin clasificar. Aunque algunos autores señalan esto como una ventaja, realmente es un problema ya que es preferible dejar áreas sin clasificar que clasificarlas sin garantías. Lo más adecuado sería cartografiar estos píxeles con problemas de clasificación para determinar cual es la razón e intentar solventarla. Normalmente serán los píxeles de borde los que den problemas de clasificación por su propia naturaleza mixta. También puede ocurrir que estos píxeles no clasificados correspondan a clases no tenidas en cuenta a la hora de definir las clases y las áreas de entrenamiento.

Por otro lado no se tiene en cuenta las desviación típicas de cada una de las bandas para cada una de las clases, así un clase con una baja desviación típica no debería absorber píxeles alejados de su centroide.

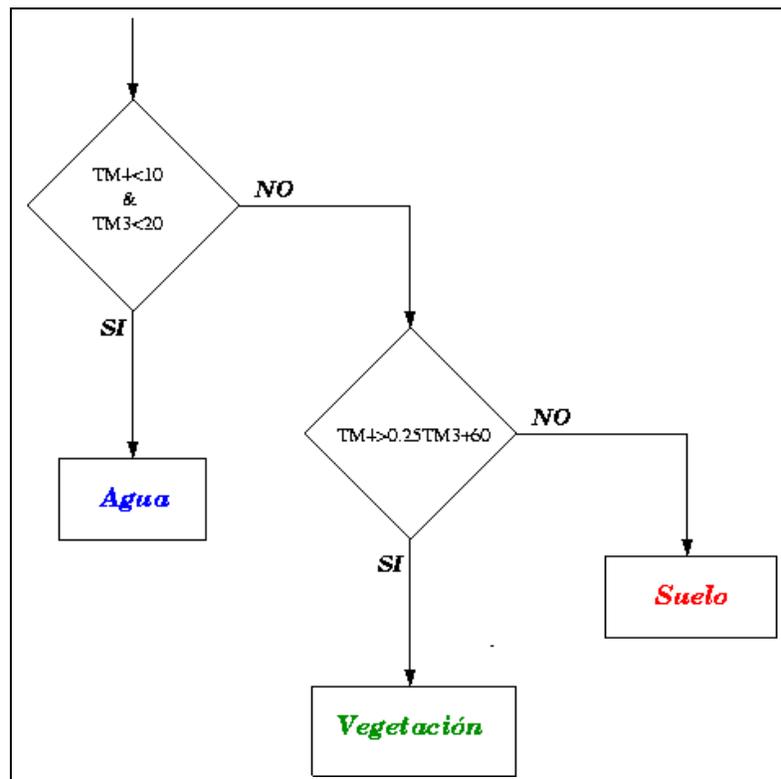


Figura 9.8: Arbol de clasificación

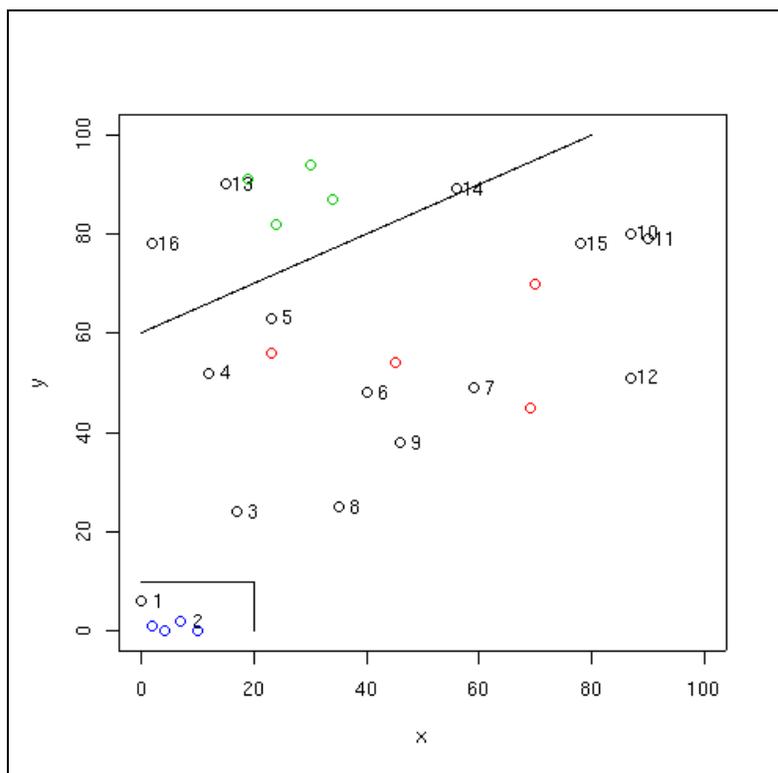


Figura 9.9: Arbol de clasificación, división del espacio de variables

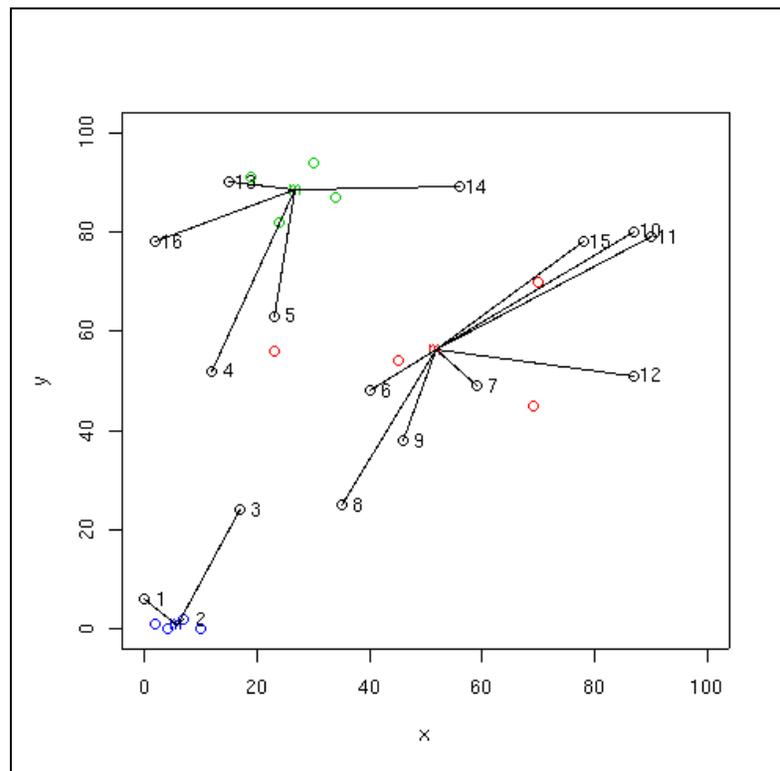


Figura 9.10: Clasificación por mínima distancia

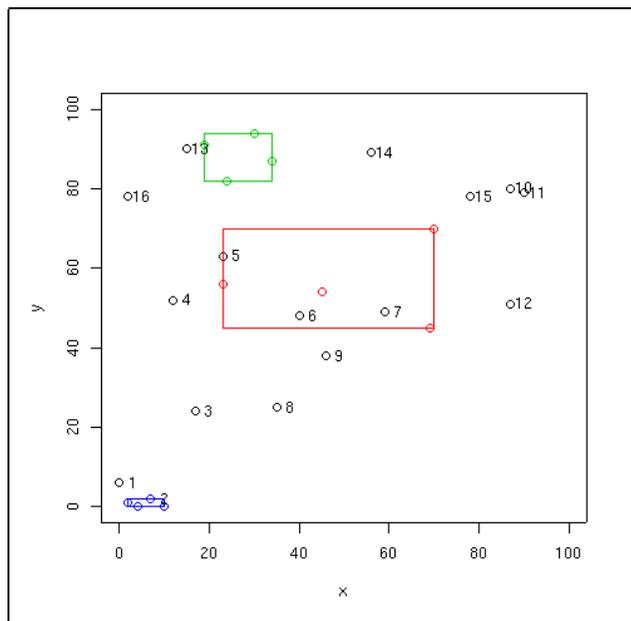


Figura 9.11: Clasificación por paralelepípedos

Por ejemplo, los píxeles de agua muestran una desviación típica reducida, es decir todos ellos muestran valores muy similares y no tendría sentido clasificar como agua un píxel (como el número 3) que se apartará de esta norma aunque estuviera más cerca de la media del agua que de cualquier otra media. Por otra parte, los píxeles de suelo muestran una alta dispersión (especialmente en cuanto a reflectividad en el rojo debido a las diferentes composiciones mineralógicas). Esta dispersión implica que píxeles bastante alejados tendrán una probabilidad relativamente alta de pertenencia a esta clase.

9.5.3 Clasificación por paralelepípedos

Un modo relativamente sencillo de solventar este problema es utilizar el método de los paralelepípedos. En este método se definen una serie de rectángulos, hiperparalelepípedos en el caso de más de 3 dimensiones, que definen las fronteras de cada clase (figura ??). La definición de cada uno de los paralelepípedos debe hacerse teniendo en cuenta los valores máximos y mínimos de reflectividad para cada una de las bandas. Con este método pueden aparecer píxeles sin asignar o píxeles asignados a varias clases.

Otro problema es que un paralelepípedo no es una forma adecuada de modelizar la dispersión de las áreas de entrenamiento, especialmente teniendo en cuenta la elevada correlación entre bandas que supone que los píxeles de las áreas de entrenamiento se dispongan como líneas oblicuas en el espacio de variables.

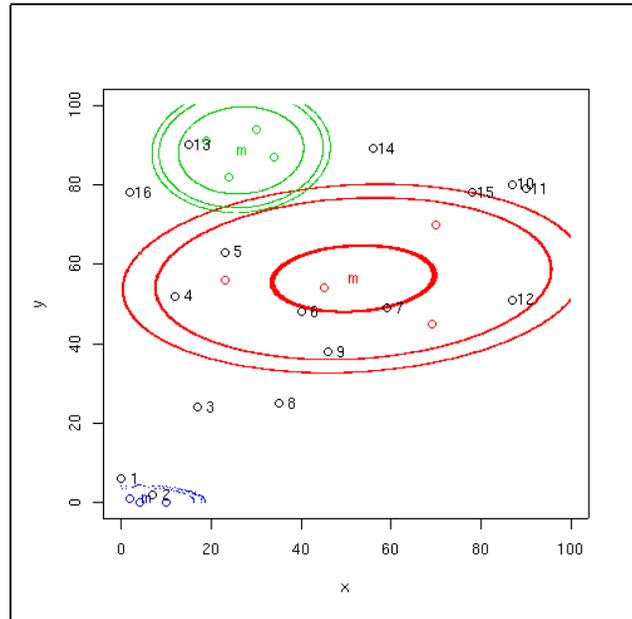


Figura 9.12: Clasificación por máxima probabilidad

9.6 Adjudicación de píxeles a clases. Técnicas estadísticas: Clasificación por máxima verosimilitud

Puesto que se tiene la media y la desviación típica de cada una de las clases, se puede utilizar algún modelo de distribución de probabilidad. El clasificador de máxima verosimilitud (o máxima probabilidad) asume que los datos siguen una función de distribución normal para asignar la probabilidad de que un pixel cualquiera pertenezca a cada una de las clases. El pixel se asigna de este modo a la clase a la que es más probable que pertenezca.

Este método puede usarse de forma automática, o puede establecerse algún criterio que permita asignar píxeles a una clase sólo si la probabilidad correspondiente es superior a determinado umbral. Permite por otro lado definir algún tipo de criterio para medir la calidad de la asignación, por ejemplo la diferencia entre la máxima probabilidad y la siguiente.

En la figura ?? aparecen una serie de elipses rodeando a cada uno de los centroides, se trata de líneas de isoprobabilidad, por tanto el pixel se clasificará como perteneciente a la clase en la que sus valores resulten más probables.

Sin embargo la hipótesis de que los datos de reflectividad siguen una distribución normal no siempre se cumple y debería verificarse.

9.7 Clasificación de imágenes. Técnicas avanzadas

Recientemente el desarrollo de los ordenadores ha permitido implementar técnicas de clasificación más sofisticadas que los métodos de mínima distancia o parelepípedos y más robustas que la clasificación de máxima verosimilitud (en el sentido de que no necesitan que los datos cumplan una determinada condición como la de seguir una distribución normal). En general si los datos cumplen estas condiciones es preferible utilizar máxima verosimilitud. El problema es que en muchos casos ni siquiera es posible saber si los datos cumplen esta condición o no.

9.7.1 Clasificación de contexto

Todos los métodos estudiados hasta ahora utilizan, para clasificar un pixel, sólo los valores de reflectividad recogidos en dicho pixel. Este hecho no es demasiado problemático si el tamaño de pixel es mayor que los elementos que los diferentes elementos físicos que componen el paisaje. Sin embargo si el tamaño del pixel es más pequeño que las unidades de paisaje podría utilizarse la información de los pixels de alrededor para estimar la pertenencia a una clase o confirmar la estimación.

Puede resultar un método útil en combinación con el de máxima probabilidad ya que permite incluir la información de los pixels circundantes para tomar una decisión en caso de que las probabilidades para dos clases sean similares o no exista una clase con probailidades de pertenencia suficientemente altas.

Contexto espacial

Se trata de incorporar otras fuentes de información distintas a las bandas para la clasificación. Entre esta información estaría la altitud, pendiente, litología, etc. Las distintas formaciones vegetales tienen mayor capacidad para desarrollarse en determinados entornos definidos en parte por estas variables. Por tanto puede ser útil establecer cual es la formación vegetal con mayor probabilidad de desarrollarse en un determinado pixel (en función de topografía y litología) y combinar esta información con la de las bandas.

9.7.2 Clasificaciones multitemporales

Si lo que se pretende es distinguir entre clases cuya respuesta espectral varían con el tiempo, por ejemplo entre diferentes usos del suelo, puede ser interesante integrar imágenes obtenidas en diferentes fechas. Imágenes de primavera y otoño tienen un gran valor discriminante ya que las superficies vegetales, especialmente las de cultivo muestran grandes diferencias.

Si se utiliza esta técnica, resulta imprescindible realizar una buena corrección atmosférica de las imágenes. Si todas son de la misma fecha, la influencia de la atmósfera es la misma para todas por tanto podemos asumir que no influye demasiado en la clasificación. Sin embargo si las imágenes proceden de diferentes fechas, la diferente influencia de la atmósfera puede enmascarar y distorsionar los cambios en la respuesta espectral.

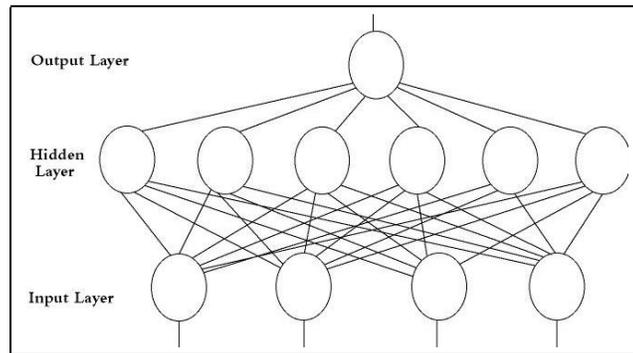


Figura 9.13: Ejemplo de red neuronal

9.7.3 Clasificación por redes neuronales

Se basan en el uso de redes neuronales artificiales que, se supone, imitan a las redes neuronales reales en el desarrollo de tareas de aprendizaje. Una neurona artificial es un objeto lógico (se trata de software no de hardware) que recibe diversas entradas, hace una suma ponderada de las mismas y produce una salida a partir de la aplicación de una función umbral a la media ponderada.

Si conectamos las salidas de unas neuronas como entradas de otras obtenemos una red neuronal. Uno de los ejemplos más típicos de red neuronal es el la Back Propagation Neural Network que aparece en la figura ??.

Consta de una capa de entrada con tantas neuronas como variables de entrada se vayan a introducir en el modelo (en el caso de la teledetección sería una por cada banda utilizada para clasificar), una capa oculta que realiza la mayor parte del cálculo y una capa de salida con tantas neuronas como posibles clases existan. En teledetección esta salida suele consistir en un valor numérico entre 0 y 1 para cada clase, cuanto mayor sea este valor más verosímil resulta que el pixel pertenezca a la clase en cuestión

Para trabajar con una red neuronal existen varias fases:

1. **Entrenamiento.** Se le introducen a la red la respuesta espectral de pixeles cuya clase se conoce y se compara la salida con la realidad. A partir de esta comparación se modifican los coeficientes de ponderación de todas las neuronas para que se obtenga la respuesta adecuada (se trata de un procedimiento automático) es decir un 1 en la clase correcta y ceros en las incorrectas
2. **Estabilización.** Al principio del entrenamiento, los factores de ponderación cambian muy deprisa, pero conforme este se desarrolla (y si las areas de entrenamiento se han seleccionado correctamente) se estabilizan (no se modifican aunque se vuelvan a introducir los pixeles de entrenamiento). En este momento finaliza la fase de entrenamiento
3. **Clasificación.** Se introducen las respuestas espectrales de los pixeles cuya clase no se conoce y se adjudican a la clase que de una respuesta más alta (que no va a ser necesariamente 1).

	A	B	C	D	E	F	Ni	n_i/N_i	n_{ii}^c	n_{ic}^i
A	50	3	0	0	2	5	60	83.3	10	21
B	4	62	3	0	0	1	70	88.5	8	10
C	4	4	70	0	8	3	86	81.4	19	6
D	0	0	0	64	0	0	64	100.0	0	3
E	3	0	2	0	71	1	78	91.0	6	10
F	10	3	1	3	0	33	50	66.0	17	10
							408	85.8	60	60

Tabla 9.1: Matriz de confusiones con 6 clases

Se trata en definitiva de un método de clasificación no paramétrico robusto que da buenos resultados cuando las respuestas espectrales de las clases no siguen una distribución normal. La clave está en el conjunto de coeficientes de ponderación que constituyen un conjunto de parámetros que deben ajustarse a unos datos de entrada y salida. Por tanto en cierto modo es equivalente a una regresión.

El conjunto de ecuaciones que queda almacenado en la red neuronal es similar a las ecuaciones que dividen el espacio de variables en varias regiones.

9.7.4 Clasificadores borrosos

Sea cual sea el procedimiento de clasificación utilizado, surge el problema de que algunos pixels resultan inclasificables, bien porque sus probabilidades de pertenencia a cualquier clase sean muy bajas (máxima verosimilitud, salidas de un método de redes neuronales, etc.) o porque aparezcan dos o más clases con muy alta probabilidad que se disputen el pixel. En estos casos tiene más sentido no clasificar los pixels de forma unívoca sino establecer cual es su posibilidad de pertenencia a cada una de las clases (el concepto de posibilidad no es exactamente igual al de probabilidad).

9.8 Validación. Error de clasificación

Existen dos posibilidades, evaluar una estimación teórica del error en función de las características del algoritmo de clasificación o analizar una serie de áreas test obtenidas del mismo modo que las áreas de entrenamiento. El segundo modo de proceder permite obtener una estimación más realista de los errores mientras la muestra de pixels para la estimación del error sea lo suficientemente grande y representativo.

Un método simple y apropiado de evaluaciones de los errores es utilizar la matriz de confusión de clases (tabla 9.1).

Con este tipo de análisis, se obtiene, no sólo una caracterización del error cometido, sino también una medida sobre la adecuación de la clasificación considerada a la realidad y de los parámetros utilizados para caracterizarlas. Puede por tanto utilizarse para definir un nuevo conjunto de clases para realizar una clasificación.

En la tabla aparece un ejemplo de matriz de confusiones, los valores en filas representan las clases reales y en columnas aparecen (en la parte izquierda de la tabla) las clases obtenidas tras la clasificación. Por tanto debe interpretarse como el número de pixels que perteneciendo a la clase fila han sido adjudicados a la clase columna.

En la parte derecha de la tabla aparece: N_i es el número total de pixels de cada clase, el porcentaje de pixels clasificados correctamente para cada clase. $(n_i)c_i$ es el número de pixels que correspondiendo a la clase i han sido adjudicados a otras, y $(n_i)c_j$ es el número de pixels adjudicados a j cuando en realidad pertenecen a otra clase. La última fila presenta la totalización de estos resultados.