

Herramientas de Teledetección Cuantitativa

Educando al clasificador

Francisco Nemiña

Unidad de Educación y Formación Masiva
Comisión Nacional de Actividades Espaciales

13 de octubre de 2016



Esquema de presentación

Escenas del capítulo anterior

Matemática

Estadística

Clasificación supervisada

Idea

Métodos

Máxima verosimilitud

Otros métodos

Práctica

Práctica



La vez pasada vimos

- ▶ Que a partir de esto podíamos definir la ρ_λ la firma espectral como una característica de cada cuerpo.
- ▶ La necesidad de definir categorías de uso y cobertura de forma concisa.
- ▶ La diferencia entre el concepto de categorías de uso y cobertura y clases espectrales.
- ▶ La importancia del espacio espectral para comprender los métodos de clasificación.
- ▶ El funcionamiento del algoritmo k-means de segmentación.



La vez pasada vimos

- ▶ Que a partir de esto podíamos definir la ρ_λ la firma espectral como una característica de cada cuerpo.
- ▶ La necesidad de definir categorías de uso y cobertura de forma concisa.
- ▶ La diferencia entre el concepto de categorías de uso y cobertura y clases espectrales.
- ▶ La importancia del espacio espectral para comprender los métodos de clasificación.
- ▶ El funcionamiento del algoritmo k-means de segmentación.



La vez pasada vimos

- ▶ Que a partir de esto podíamos definir la ρ_λ la firma espectral como una característica de cada cuerpo.
- ▶ La necesidad de definir categorías de uso y cobertura de forma concisa.
- ▶ La diferencia entre el concepto de categorías de uso y cobertura y clases espectrales.
- ▶ La importancia del espacio espectral para comprender los métodos de clasificación.
- ▶ El funcionamiento del algoritmo k-means de segmentación.



La vez pasada vimos

- ▶ Que a partir de esto podíamos definir la ρ_λ la firma espectral como una característica de cada cuerpo.
- ▶ La necesidad de definir categorías de uso y cobertura de forma concisa.
- ▶ La diferencia entre el concepto de categorías de uso y cobertura y clases espectrales.
- ▶ La importancia del espacio espectral para comprender los métodos de clasificación.
- ▶ El funcionamiento del algoritmo k-means de segmentación.



La vez pasada vimos

- ▶ Que a partir de esto podíamos definir la ρ_λ la firma espectral como una característica de cada cuerpo.
- ▶ La necesidad de definir categorías de uso y cobertura de forma concisa.
- ▶ La diferencia entre el concepto de categorías de uso y cobertura y clases espectrales.
- ▶ La importancia del espacio espectral para comprender los métodos de clasificación.
- ▶ El funcionamiento del algoritmo k-means de segmentación.



Esquema de presentación

Escenas del capítulo anterior

Matemática

Estadística

Clasificación supervisada

Idea

Métodos

Máxima verosimilitud

Otros métodos

Práctica

Práctica



Notación

Notamos a la media para la clase ω_i como

$$m_i = \frac{1}{q_i - 1} \sum_j^{q_i} x_j$$

donde q_i es la cantidad de píxeles de la clase.

Notación

La varianza como

$$\sigma_i^2 = \frac{1}{q_i - 1} \sum_j^{q_i} (x_j - m_i)^2$$

dónde los x_j pertenecen a la clase i .



Notación

Notamos a la media para la clase ω_i como

$$m_i = \frac{1}{q_i - 1} \sum_j^{q_i} x_j$$

donde q_i es la cantidad de píxeles de la clase.

Notación

La varianza como

$$\sigma_i^2 = \frac{1}{q_i - 1} \sum_j^{q_i} (x_j - m_i)^2$$

dónde los x_j pertenecen a la clase i .



Probabilidad condicional

Recordamos a la probabilidad condicional como

$$p(x|\omega_i)$$

como la probabilidad de encontrar a un píxel en el punto x del espacio espectral dado que sabemos que pertenece a la clase ω_i .



Teorema de Bayes

$$p(\omega_i|x) = \frac{p(x|\omega_i)p(\omega_i)}{p(x)}$$

Es decir, la probabilidad de que un píxel pertenezca a la clase ω_i dado que se encuentra en el punto del espacio espectral x .



Distribución de Gauss multidimensional

Si definimos a la matriz de covarianza como

$$C_i = \frac{1}{q_i - 1} \sum_j^{q_i} (x_j - m_i)(x_j - m_i)^T$$

podemos definir la distribución de Gauss en un espacio multidimensional como

$$p(x|\omega_i) \sim \exp\left(\frac{-1}{2}(x - m_i)^T C_i^{-1}(x - m_i)\right)$$



Esquema de presentación

Escenas del capítulo anterior

Matemática

Estadística

Clasificación supervisada

Idea

Métodos

Máxima verosimilitud

Otros métodos

Práctica

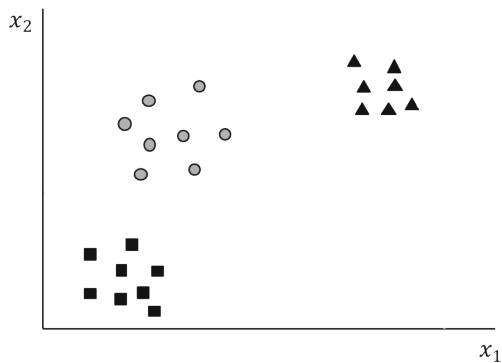
Práctica



Importante

Ahora tenemos que definir apriori cuales son las clases que queremos y como encontrarlas.

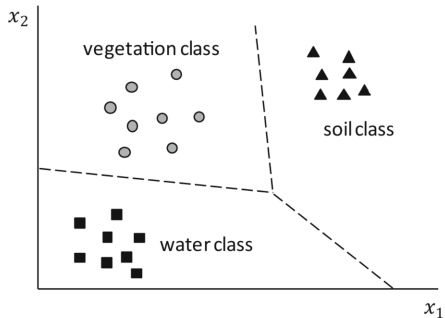




Espacio vectorial.¹

¹richards2013remote.





Clasificación del espacio vectorial a partir de clases de entrenamiento.²



Esquema general

1. Decidir cuales son las clases de intereés.
2. Elegir píxeles conocidos y representativos para cada clase a utilizar como áreas de entrenamiento.
3. Estimar los parámetros del método de clasificación.
4. Usar el clasificador para clasificar los pixeles.
5. Producir mapas temáticos para extraer información.
6. Corroborar la precisión de la clasificación con datos de campo



Esquema general

1. Decidir cuales son las clases de intereés.
2. Elegir píxeles conocidos y representativos para cada clase a utilizar como áreas de entrenamiento.
3. Estimar los parámetros del método de clasificación.
4. Usar el clasificador para clasificar los pixeles.
5. Producir mapas temáticos para extraer información.
6. Corroborar la precisión de la clasificación con datos de campo



Esquema general

1. Decidir cuales son las clases de intereés.
2. Elegir píxeles conocidos y representativos para cada clase a utilizar como áreas de entrenamiento.
- 3. Estimar los parámetros del método de clasificación.**
4. Usar el clasificador para clasificar los pixeles.
5. Producir mapas temáticos para extraer información.
6. Corroborar la precisión de la clasificación con datos de campo



Esquema general

1. Decidir cuales son las clases de intereés.
2. Elegir píxeles conocidos y representativos para cada clase a utilizar como áreas de entrenamiento.
3. Estimar los parámetros del método de clasificación.
4. Usar el clasificador para clasificar los pixeles.
5. Producir mapas temáticos para extraer información.
6. Corroborar la precisión de la clasificación con datos de campo



Esquema general

1. Decidir cuales son las clases de intereés.
2. Elegir píxeles conocidos y representativos para cada clase a utilizar como áreas de entrenamiento.
3. Estimar los parámetros del método de clasificación.
4. Usar el clasificador para clasificar los pixeles.
- 5. Producir mapas temáticos para extraer información.**
6. Corroborar la precisión de la clasificación con datos de campo



Esquema general

1. Decidir cuales son las clases de intereés.
2. Elegir píxeles conocidos y representativos para cada clase a utilizar como áreas de entrenamiento.
3. Estimar los parámetros del método de clasificación.
4. Usar el clasificador para clasificar los pixeles.
5. Producir mapas temáticos para extraer información.
6. Corroborar la precisión de la clasificación con datos de campo



Generales

- ▶ Paralelepípedos
- ▶ Distancia mínima
- ▶ **Máxima verosimilitud**
- ▶ Ángulo espectral



Clasificador Bayesiano

Si conocemos las probabilidades condicionales $p(\omega_i|x)$ entonces un píxel x pertenece a la clase ω_i si

$$p(\omega_i|x) > p(\omega_j|x)$$

si $i \neq j$.

Problema

No conocemos $p(\omega_i|x)$.



Clasificador Bayesiano

Si conocemos las probabilidades condicionales $p(\omega_i|x)$ entonces un píxel x pertenece a la clase ω_i si

$$p(\omega_i|x) > p(\omega_j|x)$$

si $i \neq j$.

Problema

No conocemos $p(\omega_i|x)$.



Solución

Usamos el teorema de Bayes y podemos escribir que un píxel x pertenece a la clase ω_i si

$$p(x|\omega_i)p(\omega_i) > p(x|\omega_j)p(\omega_j)$$

si $i \neq j$.

Función discriminante

Si definimos $g_i(x) = \log(p(x|\omega_i)p(\omega_i))$ entonces lo anterior se convierte en x pertenece a la clase ω_i si

$$g_i(x) > g_j(x)$$

si $i \neq j$.



Solución

Usamos el teorema de Bayes y podemos escribir que un píxel x pertenece a la clase ω_i si

$$p(x|\omega_i)p(\omega_i) > p(x|\omega_j)p(\omega_j)$$

si $i \neq j$.

Función discriminante

Si definimos $g_i(x) = \log(p(x|\omega_i)p(\omega_i))$ entonces lo anterior se convierte en x pertenece a la clase ω_i si

$$g_i(x) > g_j(x)$$

si $i \neq j$.



Caso Gaussiano

Si suponemos que la distribución p es normal y que, apriori la probabilidad de pertenecer a una clase es equiprobable, tenemos que

$$g_i(x) = -\log |C_i| - (x - m_i)^T C_i^{-1} (x - m_i)$$

.

Observaciones:

Como la distribución de Gauss no se anula nunca, esto puede clasificar a lo largo de todo el espacio



Caso Gaussiano

Si suponemos que la distribución p es normal y que, apriori la probabilidad de pertenecer a una clase es equiprobable, tenemos que

$$g_i(x) = -\log |C_i| - (x - m_i)^T C_i^{-1} (x - m_i)$$

.

Observaciones:

Como la distribución de Gauss no se anula nunca, esto puede clasificar a lo largo de todo el espacio



Superficies de equiprobabilidad

Si buscamos la superficies de

$$g_i = g_j$$

ese espacio queda dividido en distintos sectores donde es siempre mayor la probabilidad de pertenecer a una clase. Son

- ▶ Elipses
- ▶ Parábolas
- ▶ Hipérbolas



Superficies de equiprobabilidad

Si buscamos la superficies de

$$g_i = g_j$$

ese espacio queda dividido en distintos sectores donde es siempre mayor la probabilidad de pertenecer a una clase. Son

- ▶ Elipses
- ▶ Parábolas
- ▶ Hipérbolas



Superficies de equiprobabilidad

Si buscamos la superficies de

$$g_i = g_j$$

ese espacio queda dividido en distintos sectores donde es siempre mayor la probabilidad de pertenecer a una clase. Son

- ▶ Elipses
- ▶ Parábolas
- ▶ Hipérbolas



Superficies de equiprobabilidad

Si buscamos la superficies de

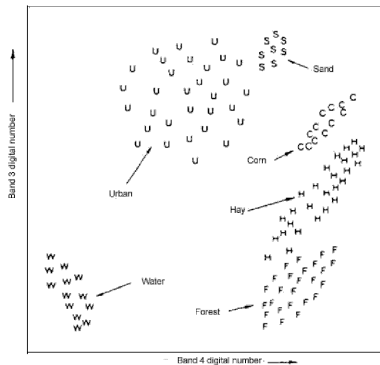
$$g_i = g_j$$

ese espacio queda dividido en distintos sectores donde es siempre mayor la probabilidad de pertenecer a una clase. Son

- ▶ Elipses
- ▶ Parábolas
- ▶ Hipérbolas



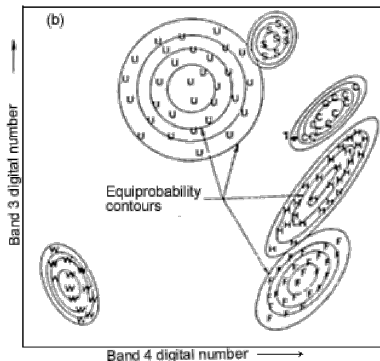
Máxima verosimilitud



Vista en el espacio vectorial.³



Máxima verosimilitud



Vista en el espacio vectorial.⁴

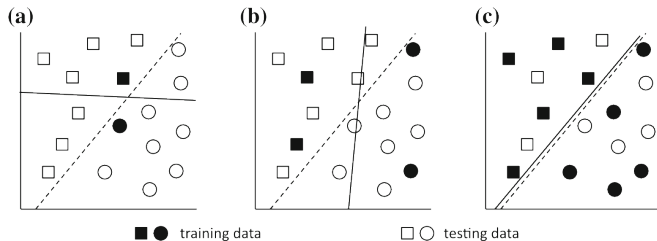


Número de píxeles necesarios

Para estimar la matriz de covarianza se necesitan al menos $N(N + 1)$ elementos. Es decir, al menos $N + 1$ píxeles.



Máxima verosimilitud



Clasificación supervisada incrementando el número de píxeles de entrenamiento.⁵

⁵richards2013remote.

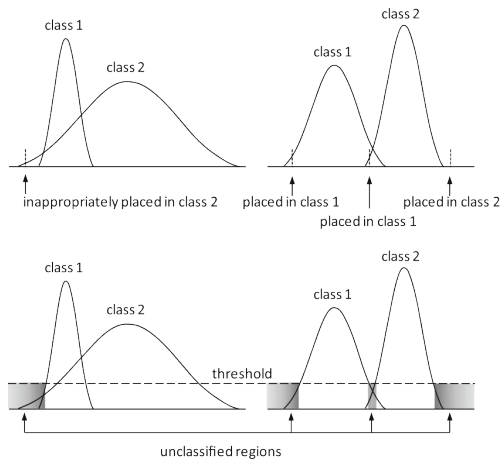


Número de píxeles necesarios

En la práctica, se necesitan entre $10N$ y $100N$ píxeles.



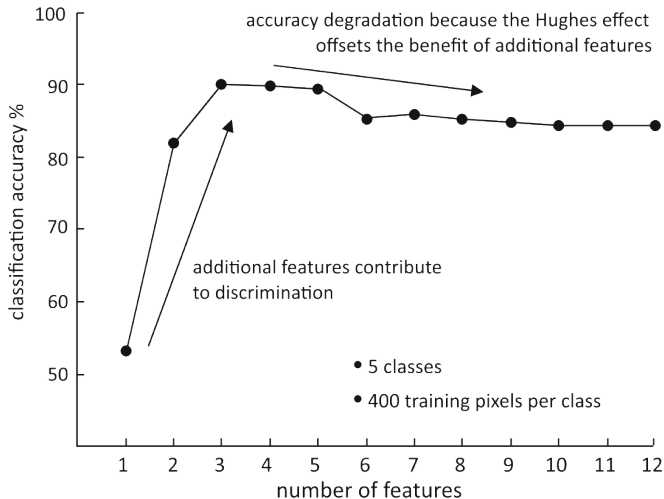
Máxima verosimilitud



Problemas de clasificación y umbral.⁶



Máxima verosimilitud



Otro problema, fenómeno de Hughes.⁷



Máxima verosimilitud

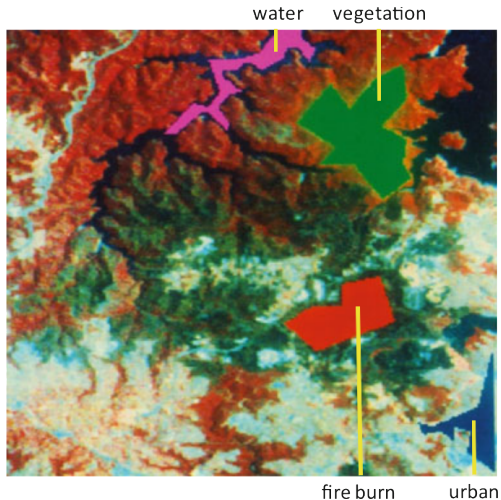


Imagen con áreas de entrenamineto.⁸

⁸richards2013remote.



Máxima verosimilitud

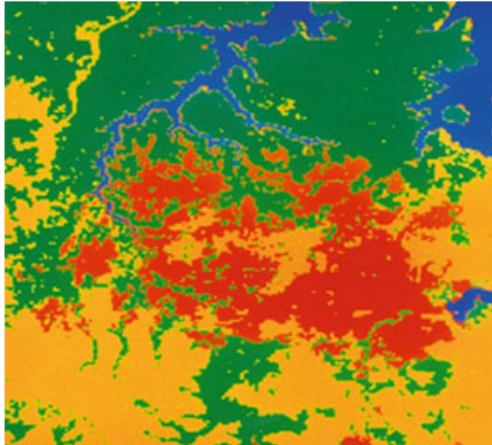


Imagen clasificada.⁹

⁹richards2013remote.



Pocos píxeles

Si contamos con pocos píxeles de entrenamiento, podemos caer en otros metodos.

- ▶ Paralelepípedos
- ▶ Distancia mínima
- ▶ Máxima verosimilitud
- ▶ Ángulo espectral

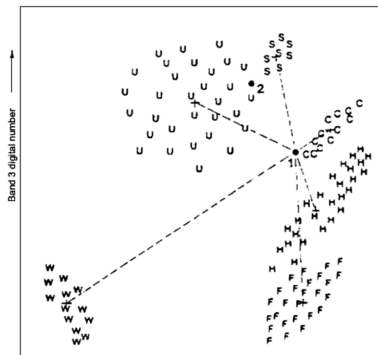


Distancia mínima

Si buscamos la superficies de $g_i = g_j$ con $g_i = 2m_i x - m_i m_i$ y me divide a mi espacio por hiperplanos.



Otros métodos



Vista en el espacio vectorial.¹⁰

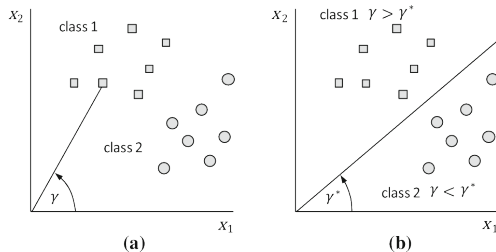


Angulo espectral

Dividimos en este caso al espacio utilizando el ángulo correspondiente a los píxeles de entrenamiento.



Otros métodos



Vista en el espacio vectorial.¹¹

¹¹richards2013remote.



Esquema de presentación

Escenas del capítulo anterior

Matemática

Estadística

Clasificación supervisada

Idea

Métodos

Máxima verosimilitud

Otros métodos

Práctica

Práctica



Actividades prácticas de la cuarta clase

1. Abra las imágenes Landsat 8 y digitalice las coberturas de interés.
2. Clasifique la imagen utilizando un vector de entrenamiento por clase.
3. Clasifique la imagen utilizando varios vectores de entrenamiento por clase.
4. Utilice la herramienta de estadísticas globales para estimar las áreas correspondientes a cada uso y cobertura.



Actividades prácticas de la cuarta clase

1. Abra las imágenes Landsat 8 y digitalice las coberturas de interés.
2. Clasifique la imagen utilizando un vector de entrenamiento por clase.
3. Clasifique la imagen utilizando varios vectores de entrenamiento por clase.
4. Utilice la herramienta de estadísticas globales para estimar las áreas correspondientes a cada uso y cobertura.



Actividades prácticas de la cuarta clase

1. Abra las imágenes Landsat 8 y digitalice las coberturas de interés.
2. Clasifique la imagen utilizando un vector de entrenamiento por clase.
3. Clasifique la imagen utilizando varios vectores de entrenamiento por clase.
4. Utilice la herramienta de estadísticas globales para estimar las áreas correspondientes a cada uso y cobertura.



Actividades prácticas de la cuarta clase

1. Abra las imágenes Landsat 8 y digitalice las coberturas de interés.
2. Clasifique la imagen utilizando un vector de entrenamiento por clase.
3. Clasifique la imagen utilizando varios vectores de entrenamiento por clase.
4. Utilice la herramienta de estadísticas globales para estimar las áreas correspondientes a cada uso y cobertura.



Esquema de presentación

Escenas del capítulo anterior

Matemática

Estadística

Clasificación supervisada

Idea

Métodos

Máxima verosimilitud

Otros métodos

Práctica

Práctica



Actividades prácticas de la primera clase

1. Clasifique la imagen por el método de máxima verosimilitud con una sola clase de entrenamiento por categoría de uso y cobertura.
2. Clasifique la imagen por el método de máxima verosimilitud con varias clases de entrenamiento por categoría de uso y cobertura.
3. Utilizar la herramienta de estadísticas globales para estimar las áreas correspondientes a cada uso y cobertura.

