



TRATAMIENTO DIGITAL DE SEÑALES

Ingeniería de Telecomunicación (4º, 2º c)

Unidad 10ª: Agrupamiento

Aníbal R. Figueiras Vidal
Jesús Cid Sueiro
Ángel Navia Vázquez

Área de Teoría de la Señal y Comunicaciones
Universidad Carlos III de Madrid

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70

amiento (“clustering”)

En ocasiones, no se puede realizar una clasificación supervisada (p.ej., los datos no están etiquetados), o no se desea hacerlo (directamente).

Puede interesar **agruparlos** según algún criterio; normalmente

para optimizar un **objetivo**

de acuerdo con una **similitud** (implica **minimizar** la similitud global) (o maximizar la similitud: ídem maximizar)

la posibilidad de satisfacer requisitos secundarios (restricciones, estructura,...)

Así se procede de forma **no supervisada**: aunque también puede decirse que hay una “supervisión integrada” según el principio seguido: cuya elección debe ser adecuada para que el agrupamiento tenga utilidad.

UCI/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70



El agrupamiento puede servir:

... como un **paso exploratorio** para una posterior clasificación:

... obteniendo información útil para ésta, p.ej., qué variables son relevantes; permitiendo proponer hipótesis o clases, e incluso una primera verificación de su validez;

... como una **reducción de datos**: pasando a trabajar con un representante de cada clase (o varios); así:

... esto reduce la dificultad de un primer diseño de un clasificador; e incluso puede ganarse en generalización;

... esto facilita un (razonable) proceso de etiquetado (por grupos): hacerlo para todas las muestras podría ser muy costoso o imposible;

... C/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

un paso para **simplificar un proceso auxiliar** en la resolución del problema de que se trate; p.ej.,

para realizar imputaciones de valores perdidos de forma local;

para reducir la carga computacional de un proceso de selección de muestras (las significativas para encontrar la solución);

de convertir un **problema global en un conjunto de problemas locales:** haciendo proponer un diseño distinto para cada grupo, o según la relación de muestra tratada con los representantes de los grupos.

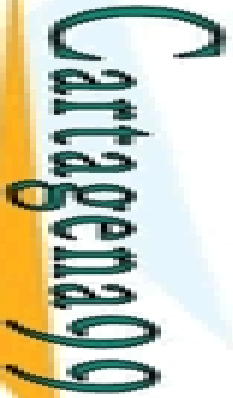
(Naturalmente, este proceso es subóptimo: ya que el agrupamiento no se hace en principio- para optimizar el subsiguiente tratamiento).

C/DTC/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP:689 45 44 70



Planteamiento para optimizar un objetivo

Los procedimientos más difundidos son los basados en considerar las clasificaciones como generadas por una mezcla de densidades:

$$p(\mathbf{x}) = \sum_{i=1}^G \text{Pr}(G_i) p(\mathbf{x} | G_i)$$

donde G es el número de grupos de la mezcla y $\{G_i\}$ dichos grupos. Se resuelven mediante la aplicación del criterio de Máxima Verosimilitud.

Posibles agrupaciones posteriores pueden hacerse determinísticamente, eligiendo el máximo, o probabilísticamente, proporcionalmente a las probabilidades (ver ejemplo de \mathbf{x}).

Otras optimizaciones, como la del error cuadrático total

$$\sum_{j=1}^G \sum_{\mathbf{x}^{(k)} \in G_j} \left\| \mathbf{x}^{(k)} - \mathbf{m}_j \right\|_2^2$$

se relacionan a planteamientos comunes con los métodos que se basan en una minimización de la similitud/disimilitud.

CI/DC/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

Agrupamiento por similitud/disimilitud

Volvamos, como ejemplo inicial, al caso de proceder con la distancia ea como medida de disimilitud.

Es obvio que:

los grupos estuviesen dados, sus representantes serían las medias grupales

$$\mathbf{m}_j = \frac{1}{\#G_j} \sum_{\forall \mathbf{x}^{(k)} \in G_j} \mathbf{x}^{(k)}$$

las medias grupales estuviesen dadas, se podría agrupar por distancias directamente

$$j^* = \arg \left[\min_j \left\{ \left\| \mathbf{x}^{(k)} - \mathbf{m}_j \right\|_2^2 \right\} \right] \Rightarrow \mathbf{x}^{(k)} \in G_{j^*}$$

hay que resolver ambas cosas a la vez.

Se hace (básicamente) mediante algoritmos secuenciales o algoritmos de bloque (bloque iterados).

C/DTC/UCIIM

Cartagena99

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70



Métodos secuenciales

El esquema general es:

1. Especializar

2. Pasar muestra a muestra

* asignando según la (di)similitud establecida

* actualizando los representantes según las asignaciones

3. Detener por resultado aceptable o saturación

En general, presentan problemas por mínimos locales.

UCIIM/DTC/C

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70
 --
 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

Algoritmo (Secuencial) Básico sigue esta forma: teniendo en cuenta cuál es la mejor variación del error cuadrático por el traslado de una muestra de un grupo a otro;

Para el grupo, G_j , al que llega la muestra, $\mathbf{x}^{(k)}$, la nueva media será:

$$\mathbf{m}'_j = \frac{1}{M_j + 1} \left(\sum_{\mathbf{x}^{(i)} \in G_j} \mathbf{x}^{(i)} + \mathbf{x}^{(k)} \right) = \frac{M_j}{M_j + 1} \mathbf{m}_j + \frac{1}{M_j + 1} \mathbf{x}^{(k)}$$

Donde M_j el número (previo) de muestras en el grupo, $\mathbf{x}^{(k)}$ la muestra incorporada, y \mathbf{m}_j la media anterior; de modo que

$$\mathbf{x}^{(i)} - \mathbf{m}'_j = \mathbf{x}^{(i)} - \mathbf{m}_j - \frac{1}{M_j + 1} (\mathbf{x}^{(k)} - \mathbf{m}_j)$$

Por lo que el error cuadrático para la muestra $\mathbf{x}^{(i)} \in G_j$ es

$$\|\mathbf{x}^{(i)} - \mathbf{m}'_j\|_2^2 = \|\mathbf{x}^{(i)} - \mathbf{m}_j\|_2^2 - 2 \frac{1}{M_j + 1} (\mathbf{x}^{(i)} - \mathbf{m}_j)^T (\mathbf{x}^{(k)} - \mathbf{m}_j) + \frac{1}{(M_j + 1)^2} \|\mathbf{x}^{(k)} - \mathbf{m}_j\|_2^2$$

Sumando para todas las $\mathbf{x}^{(i)} \in G_j$, el segundo sumando se anula

(debido a que $\sum (\mathbf{x}^{(i)} - \mathbf{m}_j) = \mathbf{0}$), y el tercero multiplicado por M_j da el incremento del error;

...
C/DTC/UCHIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70
 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

Además se incrementa el error en

$$\left\| \mathbf{x}^{(k)} - \mathbf{m}'_j \right\|_2^2 = \left\| \mathbf{x}^{(k)} - \frac{M_j}{M_j + 1} \mathbf{m}_j - \frac{1}{M_j + 1} \mathbf{x}^{(k)} \right\|_2^2 = \frac{M_j^2}{(M_j + 1)^2} \left\| \mathbf{x}^{(k)} - \mathbf{m}_j \right\|_2^2$$

Por lo que el incremento total es:

$$\left[\frac{M_j}{(M_j + 1)^2} + \frac{M_j^2}{(M_j + 1)^2} \right] \left\| \mathbf{x}^{(k)} - \mathbf{m}_j \right\|_2^2 = \frac{M_j}{M_j + 1} \left\| \mathbf{x}^{(k)} - \mathbf{m}_j \right\|_2^2$$

Análogamente, el grupo que pierde la muestra reduce su error cuadrático en

$$\frac{M_{j'}}{M_{j'} - 1} \left\| \mathbf{x}^{(k)} - \mathbf{m}_{j'} \right\|_2^2$$

Por lo que el Algoritmo (Secuencial) Básico es como se presenta a continuación.

UC/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

Algoritmo (Secuencial) Básico:

Inicializar: establecer $\{\mathbf{x}^{(k)} \in G_j\}$ y fijar $\mathbf{m}_j = \text{promedio}\{\mathbf{x}^{(k)} \in G_j\}$

Tomar una muestra $\mathbf{x}^{(k)}$; sea

si $M_j = 1$, volver a 2;

si $M_j > 1$, calcular

$$\text{decr}_{j'} = \frac{M_{j'}}{M_{j'} - 1} \|\mathbf{x}^{(k)} - \mathbf{m}_{j'}\|_2^2 \qquad \text{incr}_j = \frac{M_j}{M_j + 1} \|\mathbf{x}^{(k)} - \mathbf{m}_j\|_2^2, \quad j \neq j'$$

Si incr_{j^*} es mínimo, asignar a G_{j^*} la muestra y recalcular las medias:

$$\mathbf{m}_{j^*} \rightarrow \frac{M_{j^*}}{M_{j^*} + 1} \mathbf{m}_{j^*} + \frac{1}{M_{j^*} + 1} \mathbf{x}^{(k)}$$

$$\mathbf{m}_{j'} \rightarrow \frac{M_{j'}}{M_{j'} - 1} \mathbf{m}_{j'} - \frac{1}{M_{j'} - 1} \mathbf{x}^{(k)}$$

Volver a 2 hasta que se cumpla el criterio de parada

La asignación inicial puede hacerse mediante una etapa del algoritmo K-means, que veremos a continuación. Una asignación completamente aleatoria puede plantear problemas de mínimos locales.

CI/DTIC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

ritmos paralelo

en como esquema general:

icializar

asar muestra a muestra la totalidad de ellas

asignando según la (di)similitud establecida

tualizar los representantes según las asignaciones anteriores

etener por resultado aceptable o saturación.

Naturalmente, también presentan problemas por mínimos locales.

El bien conocido **Algoritmo K-medias** (C-medias, ISODATA básico)

signa por comparación directa de distancias. Procede como se indica a

ntinuación.

C/DTC/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

--

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70

ritmo K-medias:

inicializar: fijar $\{\mathbf{m}_j\}$

para todas las muestras $\mathbf{x}^{(k)}$

$$j^* = \arg \left[\min_j \left\{ \left\| \mathbf{x}^{(k)} - \mathbf{m}_j \right\|_2^2 \right\} \right] \Rightarrow \mathbf{x}^{(k)} \in G_{j^*}$$

recalcular las medias:

$$\mathbf{m}_j = \frac{1}{M_j} \sum_{\mathbf{x}^{(k)} \in G_j} \mathbf{x}^{(k)}$$

Volver a 2 hasta que se cumpla el criterio de parada.

C/DTC/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

Sobre el número de grupos

Los algoritmos anteriores lo dan por supuesto: naturalmente, no tiene por qué ser así.

Puede determinarse extendiendo los algoritmos a:

Formas **crecientes**: aumentando el número de grupos por **división** de los existentes o por **creación** de grupos nuevos;

Formas **autoconstructivas**: completando el proceso de crecimiento por **división** de grupos existentes.

No son aconsejables las formas decrecientes, por motivos computacionales).

siguen algunos ejemplos clásicos.

C/DTC/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70

Algoritmo LBG (Linde-Buzo-Gray)

Definir $G=2$

Ejecutar el Algoritmo K-medias

Mientras no se alcance el criterio de parada (resultado aceptable o saturación o $G_{\text{máx}}$), duplicar G y volver a 2.

Duplica para poder indexar los grupos en binario).

Un modo habitual de duplicar es dividir las medias del modo:

$$\mathbf{m}_j^+ = \mathbf{m}_j + \alpha \sigma_j \mathbf{u}$$

$$\mathbf{m}_j^- = \mathbf{m}_j - \alpha \sigma_j \mathbf{u}$$

\mathbf{u} es un vector unitario aleatorio, σ_j es la desviación típica del grupo, y α es un parámetro seleccionable entre 0 y 1.

UCI/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70



K-MEANS (Iterative Self-Organizing Data Analysis Technique Algorithm)

Es un algoritmo que incluye (versión “standard”):

1. Selección de grupos: a la que se procede cuando

la varianza muestral rebasa un cierto umbral, y

la distancia promedio de sus muestras al representante es la mayor de todas las que se han considerado, o cuando el número actual de grupos es menor que un cierto valor (normalmente, la mitad del número deseado o esperado)

2. Fusión de grupos: fundiendo dos si su distancia entre representantes es menor que un cierto umbral; procediendo (en caso de varias candidaturas) en orden de prioridades crecientes y hasta un número máximo de fusiones en cada paso.

UCI/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70

Algoritmo de Hall

Es un algoritmo secuencial que incluye la creación de grupos mediante el uso de un umbral θ :

1. Se crea la primera clase con una muestra, $(\mathbf{x}^{(1)})$;

2. Se presenta la muestra $\mathbf{x}^{(k)}$ y se elige el grupo G_j más cercano:

$$\text{si } \|\mathbf{x}^{(k)} - \mathbf{m}_j\|_2^2 < \theta \quad \text{o si } G = G_{\text{máx}}, \quad \mathbf{x}^{(k)} \in G_j$$

$$\text{si } \|\mathbf{x}^{(k)} - \mathbf{m}_j\|_2^2 > \theta \quad \text{y } G < G_{\text{máx}}, \quad \text{se crea una nueva clase con } \mathbf{x}^{(k)}$$

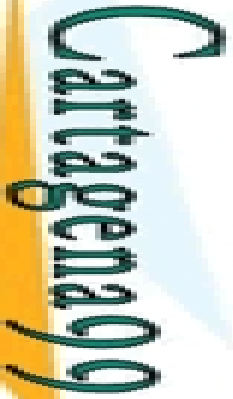
3. Si es necesario, se actualizan las representaciones de los grupos;

4. Se termina al terminarse las muestras.

UCI/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70



Obviamente, es defecto de este algoritmo la asignación de muestras a s antes de que éstos estén totalmente definidos.

Versión **Modificada**: para reducir el problema,

- parte de las muestras se utilizan como se ha visto;
- el resto se asigna directamente.

Otro inconveniente (común con todos los secuenciales) es la dependencia den de presentación de las muestras.

Versión con **Dos Umbrales**: para reducir la dependencia,

- para asignar a un grupo se utiliza θ_1
- para crear un nuevo grupo, $\theta_2 > \theta_1$

os casos intermedios, se **demora** la asignación de la muestra a una etapa ior.

C/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70
 --
 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70

El parámetro θ (o el par θ_1, θ_2) tiene una importante utilidad: como $G = G(\theta)$, se pueden explorar resultados para varios valores de θ , con ordenaciones de los datos para cada valor; eligiendo el G más frecuente para cada θ , se puede adoptar como θ más conveniente el valor que esté en el centro de la zona plana más extensa (elección para mayor robustez).

C/DTC/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

-- --

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP: 689 45 44 70



Todos los algoritmos de error cuadrático tienen versiones generalizadas iatas: en que la distancia euclídea se sustituye por una medida (general) disimilitud $d(\mathbf{x}^{(k)}, \mathbf{c}_j)$, en que \mathbf{c}_j es un **centroide**

$$\mathbf{c}_j : \min_{\mathbf{c}_j} \sum_{\mathbf{x}^{(k)} \in G_j} d(\mathbf{x}^{(k)}, \mathbf{c}_j)$$

...

amente se agrupa por mínima disimilitud; minimizando así la disimilitud

$$D = \sum_{j=1}^G \sum_{\mathbf{x}^{(k)} \in G_j} d(\mathbf{x}^{(k)}, \mathbf{c}_j)$$

C/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

Si d es una medida de **distorsión**, se obtiene lo que se llama un **cuantificador Vectorial** (VQ, “Vector Quantizer”): representar las \mathbf{x}_i por sus \mathbf{c}_j , con un buen diseño, lo mismo para muestras \mathbf{x} de la población representada por $\{\mathbf{x}_i\}$ da lugar a una **distorsión** (media) mínima; con la ventaja de que, a efectos de transmisión, los \mathbf{c}_j se pueden sustituir por índices binarios (véase la duplicación del Algoritmo LBG). Los algoritmos paralelos aplicados a la codificación (K-medias modificado, LBG modificado) se suelen denominar **algoritmos de Lloyd Generalizados** (Lloyd fue el diseñador del cuantificador óptimo).

Objetivo: Influencia de los errores de canal en los VQ; codificación fuente más

C/DTC/UCIIM

Cartagena99

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70



Agrupamiento mixto

Puede ser interesante agrupar según una similitud, pero elegir como representante otro distinto del centroide correspondiente a ésta: p.ej., si se elige el agrupamiento como vía de reducción de datos, puede convenir separar cosas.

Un ejemplo relevante es el **Algoritmo de Kohonen**, que consiste en actualizar el centroide de acuerdo con el producto escalar y actualizar por acercamiento

C/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70
 --
 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70

Es decir: los representantes $\{\mathbf{w}_j\}$ compiten para quedarse con las ras $\{\mathbf{x}^{(k)}\}$ de acuerdo con

$$j^* = \arg \left[\max_j \left\{ \mathbf{w}_j^T \mathbf{x}^{(k)} \right\} \right] \Rightarrow \mathbf{x}^{(k)} \in G_{j^*}$$

Actualización se hace según

$$\mathbf{w}_j(k+1) = \begin{cases} \mathbf{w}_j(k) + \eta(k) [\mathbf{x}^{(k)} - \mathbf{w}_j(k)], & j = j^* \\ \mathbf{w}_j(k), & j \neq j^* \end{cases}$$

se llama modo **acretivo** o WTA: “Winner Takes All”; los modos **olativos** actualizan también representantes perdedores, en menor grado).

$\eta(k)$ empieza con un valor cercano a 1 y también se hace tender a 0.

C/DTC/UCHIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70



El tipo de actualización (aprendizaje) que aquí aparece por primera vez, ma no supervisada, se denomina **Hebbiano: refuerza la capacidad de** (cualitativamente, en este caso) del representante vencedor ante casos antes (nótese que, al incluir $\eta \mathbf{x}^{(k)}$ en $\mathbf{w}_j(k+1)$, el producto escalar de éste muestras que se parezcan a $\mathbf{x}^{(k)}$ tenderá a crecer).

Obviamente, esta forma de agrupamiento no minimiza el error tico: ya que la victoria se produce según un producto escalar, y, aunque "equilibra" la competición, los grupos establecidos no lo serán para izar el error cuadrático (salvo que todas las muestras tengan igual o).

C/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70
 --
 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70

Como, obviamente, dentro de cada grupo el representante \mathbf{w}_j toma como al acabar el entrenamiento la media de las muestras del grupo (ya que va orando valores de las muestras capturadas y olvidando su propio pasado), se minimiza es el error cuadrático para una asignación dada; es decir, en

$$\sum_{\forall \mathbf{x}^{(k)}} \sum_{j=1}^G \text{ind}_j(\mathbf{x}^{(k)}) \|\mathbf{x}^{(k)} - \mathbf{w}_j\|_2^2$$

$\text{ind}_j(\mathbf{x}^{(k)})$ el indicador de pertenencia de $\mathbf{x}^{(k)}$ a G_j (grupo de \mathbf{w}_j) : con el amiento acretivo visto, vale 1 ó 0.

Si se desea un entrenamiento global, pueden utilizarse indicadores de encia suavizados (derivables) $\{z_j(\mathbf{x}^{(k)})\}$ (con suma unitaria), lo que le a un funcionamiento interpolativo (p.ej., según el valor de productos res); y se puede aplicar gradiente para el entrenamiento

$$\mathbf{w}_j^{(k+1)} = \mathbf{w}_j^{(k)} + \eta(k) \frac{\partial \left(\left\{ z_j(\mathbf{x}^{(k)}) \right\} \|\mathbf{x}^{(k)} - \mathbf{w}_j^{(k)}\|_2^2 \right)}{\partial \mathbf{w}_j^{(k)}}, \forall j$$

C/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70



ización

Se ha mencionado que los algoritmos que preceden son sensibles a mínimos locales. La sensibilidad es alta por tratarse de algoritmos de carácter **competitivo**: los representantes se disputan las muestras, y se quedan a ellas cuando ganan; pues bien; cualquier desequilibrio en la actualización llevará a una solución no deseada.

Evidentemente, tales desequilibrios se darán con facilidad si la actualización es pobre.

Utilizar como representantes iniciales G muestras tomadas al azar suele ser suficiente para los algoritmos paralelo, ya que la actualización de los representantes tiene lugar después de una competición entre todas las muestras.

C/DTC/UCIIM

CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP:689 45 44 70

En el caso de los algoritmos secuenciales, la dificultad es mayor: un agente que gane y se acerque a una zona con muchas muestras puede "dominarla", y reducir la actividad de los otros representantes. Para evitarlo, hay que "equilibrar" la competición en su fase inicial; pudiendo emplearse:

Métodos tradicionales de **modificación** de muestras, como añadir a las muestras un ruido (que las dispersa) de nivel decreciente;

o utilizar el método de **refuerzo radial**, trabajando con

$$\beta(k) \mathbf{x}^{(k)}$$

donde $\beta(k)$ es una función que comienza en $\beta(0)$ bajo y creciente hacia 1 con k ;

pero estos métodos tienen efectos correctores limitados y ralentizan el agrupamiento;

IC/DTC/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
LLAMA O ENVÍA WHATSAPP: 689 45 44 70

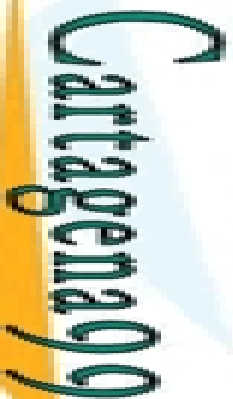
ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
CALL OR WHATSAPP: 689 45 44 70

Métodos de arrastre, que acercan a la muestra en cuestión tanto el **centroide ganador** como los de un entorno alrededor de él, éstos con pasos **decrecientes** a lo largo del proceso (también puede disminuir el entorno);

Métodos de penalización, que dificultan que los centroides que ganan **sigan** lo sigan haciendo; como los procedimientos “**sensibles a la frecuencia**” (FS, “Frequency Sensitive”), que clasifican de acuerdo con la **frecuencia** multiplicada por el número previo de victorias de cada centroide (o **potencia** creciente de este número).

Los dos últimos tipos ofrecen buenas prestaciones.

C/DTC/UCIIM



CLASES PARTICULARES, TUTORÍAS TÉCNICAS ONLINE
 LLAMA O ENVÍA WHATSAPP: 689 45 44 70

 ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS
 CALL OR WHATSAPP: 689 45 44 70