

CAPÍTULO

4. MEMORIAS ASOCIATIVAS GEOMÉTRICAS

En este capítulo se establecen las bases para el desarrollo del nuevo modelo de memorias asociativas, las llamadas memorias asociativas geométricas. Mientras que los primeros modelos de memorias asociativas utilizaban operadores del álgebra convencional para su funcionamiento y las memorias morfológicas basaban su operación en la morfología matemática, el modelo propuesto en esta tesis utiliza, para su operación, operadores del álgebra geométrica; en particular se utiliza el modelo conforme.

Se desarrollaron dos modos de operación de las memorias asociativas geométricas que atacan, precisamente, dos de los problemas más importantes del reconocimiento de patrones: el problema de la clasificación y el problema de la restauración.

Su función, como se verá más adelante, es clasificar un patrón si y solo si el patrón se encuentra dentro de vecindad de soporte (híper-esfera) de esa clase.

4.1 Modo para clasificación de patrones

En las siguientes secciones se explican tres métodos para construir vecindades híper-esféricas de forma no iterativa.

4.1.1 Creación de las vecindades híper-esféricas

La construcción de las vecindades híper-esféricas implica, primero encontrar el centro de cada híper-esfera y después calcular el radio adecuado. En la literatura se han reportado algunos procedimientos para lograr esto con álgebra geométrica conforme. En las siguientes líneas se describen tres de ellos.

El primer método se describe en (Cruz, Barrón y Sossa 2009), donde dado un conjunto de puntos $\mathbf{p} = \mathbf{p}^i$ $_{i=1}^m$ se construye una vecindad híper-esférica. El centro se calcula de una manera muy sencilla:

$$\mathbf{c} = \sum_{i=1}^m \mathbf{p}_i / m. \quad (1.1)$$

En otras palabras, el centro de cada híper-esfera es el promedio de todos los puntos. La expresión (4.2) se utiliza para calcular el radio.

$$-\frac{1}{2} \gamma^2 = \min \left[C \cdot P^i \right], \forall_{i=1, \dots, m}, \quad (1.2)$$

donde C y P^i son la representación conforme de \mathbf{c} y \mathbf{p}^i respectivamente. El procedimiento garantiza que todos los patrones en una clase estarán cubiertos por una híper-esfera. Como el centro y el radio se calculan en diferentes etapas, la híper-esfera resultante podría no estar balanceada, como se observa en la Figura 1A.

Otra desventaja importante de este procedimiento es su alto costo computacional. En (Hildenbrand 2005) se presenta un método para ajustar conjuntos de puntos en híper-planos o híper-esferas usando una aproximación por mínimos cuadrados. En ese trabajo se describe el ajuste de híper-esferas (Figura 1B). Ese algoritmo usa una medida de distancia entre puntos e híper-esferas con ayuda del producto interior. Utiliza una aproximación por mínimos cuadrados para minimizar el cuadrado de la distancia entre un punto y una híper-esfera.

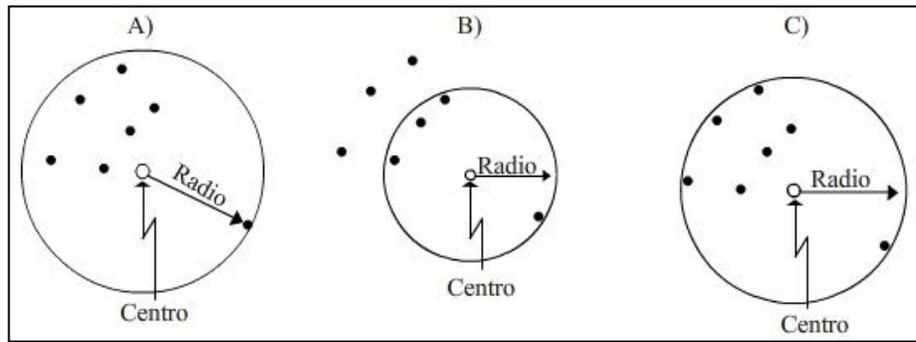


Figura 1. Tres soluciones gráficas para encontrar el centro y el radio de una hiper-esfera.

Con este procedimiento es posible crear vecindades hiper-esféricas que se ajusten a un conjunto de patrones. Un inconveniente que tiene este método es que, algunas veces, algunos puntos pueden aparecer afuera de la resultante; por lo que los patrones usados en el entrenamiento no serán clasificados correctamente.

En (Barrón, Cruz, y otros 2008) se presentó un método para construir de manera óptima una envoltura convexa (del inglés *convex hull*) esférica usando programación cuadrática y álgebra geométrica. Una envoltura convexa es importante porque es una forma más natural de modelar una clase con sus patrones correspondientes y, por lo tanto, clasificarlos de acuerdo a esta clase. El método combina las características de los métodos mencionados en los párrafos previos. El método, además, ajusta una hiper-esfera óptima pero conteniendo todos los puntos (Figura 1C).

Los métodos mencionados se pueden usar para construir una vecindad hiper-esférica para una clase específica usando los puntos (patrones) de esa misma clase. Pero no toman en cuenta los patrones de las otras clases o la separación de las mismas.

En la siguiente sección se presenta un nuevo método para construir vecindades de decisión hiper-esféricas, pero tomando en cuenta los puntos que estarán dentro de la vecindad y los puntos que deben estar afuera de la misma.

4.1.2 Nuevo método para crear vecindades esféricas

Antes de explicar el método se enunciará una nueva definición, partiendo de la **¡Error! No se encuentra el origen de la referencia.** sobre separabilidad lineal:

Definición 1. Cuando dos conjuntos de puntos en \mathbb{R}^n son completamente separados por una hiper-esfera, se dice que estos conjuntos son esféricamente separables.

En este caso la separación consiste en tener un conjunto de puntos *dentro* de la hiper-esfera y el otro conjunto de puntos *afuera* de la hiper-esfera. A partir de esta definición se enunciará el siguiente teorema que tiene que ver con la relación entre la separabilidad esférica y la lineal.

Teorema 1. Si dos conjuntos de puntos en \mathbb{R}^n son linealmente separables, entonces estos conjuntos son también esféricamente separables.

Demostración. Sean dos conjuntos de puntos linealmente separables en \mathbb{R}^n . De la definición 1 existe un hiper-plano T que los separa. Pero de **¡Error! No se encuentra el origen de la referencia.** y **¡Error! No se encuentra el origen de la referencia.** el hiper-plano T es una hiper-esfera que pasa por el punto al infinito. Entonces, también existe una hiper-esfera que los separa, por lo tanto ambos conjuntos son también esféricamente separables ■

Hay que notar que el Teorema 1 no garantiza que dos conjuntos de puntos esféricamente separables sean también linealmente separables.

A continuación se describe una nueva propuesta, inspirada en ideas de (Barrón, Cruz, y otros 2008). Ésta sirve para encontrar una vecindad hiper-esférica que toma en cuenta los patrones de la clase dentro de la hiper-esfera y los patrones de las otras clases.

Sean p^i $_{i=1}^l$ y p^j $_{j=l+1}^m$ dos conjuntos de puntos esféricamente separables en \mathbb{R}^n , tal que p^i son puntos que pertenecen a una clase y p^j son puntos que pertenecen a otra clase. El problema es encontrar una hiper-esfera S óptima (en términos de programación cuadrática) con el menor error cuadrático, donde los puntos p^i estén dentro de S y los de p^j estén afuera. En otras palabras, resolver:

$$\arg \min_S \sum_{i=1}^m P^i \cdot S^2. \quad (1.3)$$

Sujeto a (4.4) para los puntos dentro de la hiper-esfera y a (4.5) para los puntos fuera de ella.

$$P^i \cdot S \geq 0, i = 1, \dots, l. \quad (1.4)$$

$$P^j \cdot S < 0, j = l+1, \dots, m. \quad (1.5)$$

Para poder encontrar una hiper-esfera óptima se debe utilizar un algoritmo de optimización; así que hay que cambiar la expresión (1.3) para ponerla en notación euclidiana y en términos de algún problema clásico de optimización.

Partiendo de (1.3) y considerando que las hiper-esferas están en forma canónica, tales que el término S_{n+2} de las mismas se puede eliminar, se obtienen:

$$\begin{aligned} \sum_{i=1}^m P^i \cdot S^2 &= \sum_{i=1}^m \left(\mathbf{p}^i \cdot \mathbf{c} - S_{n+1} - \frac{1}{2} \mathbf{p}^{i \cdot 2} \right)^2 \\ &= \sum_{i=1}^m \left([\mathbf{p}^i \cdot \mathbf{c} - S_{n+1}] - \frac{1}{2} \mathbf{p}^{i \cdot 2} \right)^2 \\ &= \sum_{i=1}^m \left([\mathbf{p}^i \cdot \mathbf{c} - S_{n+1}]^2 - [\mathbf{p}^i \cdot \mathbf{c} - S_{n+1}] \mathbf{p}^{i \cdot 2} + \frac{1}{4} \mathbf{p}^{i \cdot 4} \right), \end{aligned} \quad (1.6)$$

hay que recordar que $S_{n+1} = \frac{1}{2} \mathbf{p}^{i \cdot 2} - \gamma^2$, siendo γ el radio de la hiper-esfera. Con lo que la expresión (1.3) se puede ver como:

$$\begin{aligned} \sum_{i=1}^m P^i \cdot S^2 &= \sum_{i=1}^m \mathbf{p}^i \cdot \mathbf{c} - S_{n+1}^2 \\ &\quad + \sum_{i=1}^m -\mathbf{p}^i \cdot \mathbf{c} + S_{n+1} \mathbf{p}^{i \cdot 2} + \frac{1}{4} \sum_{i=1}^m \mathbf{p}^{i \cdot 4}. \end{aligned} \quad (1.7)$$

El tercer término de la expresión es irrelevante porque no depende del parámetro S y se puede omitir. Sin pérdida de generalidad esta expresión puede ser escrita en notación matricial euclidiana como en (4.8), donde \mathbf{W} y \mathbf{F} son matrices, cuyas

componentes se observan en (4.9) y en (4.10) respectivamente, y $\mathbf{x} = [S_1, S_2, \dots, S_{n+1}]$. Además, hay que recordar que $S_k = c_k \forall_{k=1, \dots, n}$.

$$\sum_{i=1}^m \left(\sum_{k=1}^{n+1} \mathbf{W}_{i,k} \mathbf{x}_k \right)^2 + \sum_{i=1}^m \left(\sum_{k=1}^{n+1} \mathbf{F}_{i,k} \mathbf{x}_k \right). \quad (1.8)$$

$$\mathbf{W}_{i,k} = \begin{cases} p_k^i & \text{para } k = 1, \dots, n \\ -1 & \text{para } k = n+1 \end{cases} \quad (1.9)$$

$$\mathbf{F}_{i,k} = \begin{cases} -p_k^i & p^{i-2} & \text{para } k = 1, \dots, n \\ p^{i-2} & & \text{para } k = n+1 \end{cases}. \quad (1.10)$$

Sea $\mathbf{w}^i = [\mathbf{W}_{i,1}, \mathbf{W}_{i,2}, \dots, \mathbf{W}_{i,n+1}]$, entonces para el término de la izquierda de la expresión (1.8) y considerando $\mathbf{w}^{i \ t}$ que es el transpuesto del vector \mathbf{w}^i , se tiene:

$$\begin{aligned} \sum_{i=1}^m \left(\sum_{k=1}^{n+1} \mathbf{W}_{i,k} S_k \right)^2 &= \sum_{i=1}^m \mathbf{w}^{i \ t} \mathbf{x}^2 \\ &= \mathbf{w}^{1 \ t} \mathbf{x}^2 + \dots + \mathbf{w}^{m \ t} \mathbf{x}^2 \\ &= \mathbf{w}^{1 \ t} \mathbf{x} \mathbf{w}^{1 \ t} \mathbf{x} + \dots + \mathbf{w}^{m \ t} \mathbf{x} \mathbf{w}^{m \ t} \mathbf{x} \\ &= \mathbf{w}^{1 \ t} \mathbf{x} \mathbf{w}^{1 \ t} + \dots + \mathbf{w}^{m \ t} \mathbf{x} \mathbf{w}^{m \ t} \mathbf{x} \\ &= [\mathbf{x}^t \mathbf{w}^1 + \dots + \mathbf{x}^t \mathbf{w}^m] \mathbf{W} \mathbf{x} \\ &= \mathbf{x}^t \mathbf{W}^t \mathbf{W} \mathbf{x}. \end{aligned} \quad (1.11)$$

Ahora, sea $\mathbf{H} = \mathbf{W}^t \mathbf{W}$ entonces se obtiene la siguiente expresión:

$$\sum_{i=1}^m \left(\sum_{k=1}^{n+1} \mathbf{W}_{i,k} S_k \right)^2 = \mathbf{x}^t \mathbf{H} \mathbf{x}. \quad (1.12)$$

Para el término de la derecha de la expresión (1.8), sea $\mathbf{y}_k = \sum_{i=1}^m \mathbf{F}_{i,k}$ entonces:

$$\begin{aligned} \sum_{i=1}^m \left(\sum_{k=1}^{n+1} \mathbf{F}_{i,k} \mathbf{x}_k \right) &= \sum_{k=1}^{n+1} \left(\sum_{i=1}^m \mathbf{F}_{i,k} \mathbf{x}_k \right) \\ &= \sum_{k=1}^{n+1} \mathbf{y}_k \mathbf{x}_k . \end{aligned} \quad (1.13)$$

Si $\mathbf{y} = [\mathbf{y}_1, \dots, \mathbf{y}_{n+1}]$ se obtiene la siguiente expresión:

$$\sum_{i=1}^m \left(\sum_{k=1}^{n+1} \mathbf{F}_{i,k} \mathbf{S}_k \right) = \mathbf{y}^t \mathbf{x}. \quad (1.14)$$

Al usar las expresiones (1.12) y (1.14), la expresión (1.3) se ha convertido a una notación matricial euclidiana:

$$\mathbf{x}^t \mathbf{H} \mathbf{x} + \mathbf{y}^t \mathbf{x}. \quad (1.15)$$

La restricción (1.4) para puntos dentro de la hiper-esfera se modifica de la siguiente forma:

$$\begin{aligned} P^i \cdot S &\geq 0 \\ \mathbf{p}^i \cdot \mathbf{c} - S_{n+1} - \frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i &\geq 0 \\ \mathbf{p}^i \cdot \mathbf{c} - S_{n+1} &\geq \frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i \\ \sum_{k=1}^{n+1} -\mathbf{W}_{i,k} S_k &\leq -\frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i , \end{aligned} \quad (1.16)$$

donde $\mathbf{W}_{i,k}$ se definió en (1.9), la expresión (1.16) se puede escribir como:

$$-\mathbf{W} \mathbf{x}^t \leq -\frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i , \quad (1.17)$$

donde \mathbf{x} se definió anteriormente. La restricción (1.5) para los puntos afuera de la hiper-esfera se modifica a:

$$\begin{aligned}
p^j \cdot S &< 0 \\
p^j \cdot c - S_{n+1} - \frac{1}{2} p^{j^2} &< 0 \\
p^j \cdot c - S_{n+1} &< \frac{1}{2} p^{j^2} \\
\sum_{k=1}^{n+1} \mathbf{W}_{j,k} S_k &< \frac{1}{2} p^{j^2}.
\end{aligned} \tag{1.18}$$

Esta expresión puede re-escribirse como:

$$\mathbf{W}x^t < \frac{1}{2} p^{j^2}. \tag{1.19}$$

Sea $\mathbf{A} = -\mathbf{W}, \mathbf{W}$ y sea \mathbf{b} un vector cuya i -ésima componente es $-\frac{1}{2} p^{i^2}$ para $i=1, \dots, l$ y $\frac{1}{2} p^{i^2} - \varepsilon$ para $i=l+1, \dots, m$, la constante ε es una cantidad positiva muy pequeña que se utiliza para cambiar el “<” de la desigualdad (1.19) a un “≤”. Las restricciones (1.17) y (1.19) son entonces convertidas a una notación matricial euclidiana del tipo:

$$\mathbf{A}x \leq \mathbf{b}. \tag{1.20}$$

Finalmente, la expresión (1.3) se cambia a un problema clásico de optimización con restricciones:

$$\begin{aligned}
\min_x \quad & x^t \mathbf{H}x + y^t x \\
\text{s.a.} \quad & \mathbf{A}x \leq \mathbf{b}.
\end{aligned} \tag{1.21}$$

La hiper-esfera óptima S se encuentra al resolver la expresión (1.21), donde $S_k = x_k$ para $k=1, \dots, n+1$ y $S_{n+2} = 1$. Está claro que al incluir también en las restricciones los puntos que se encuentran fuera de la hiper-esfera, la solución S es una superficie de

separación que permite diferenciar entre dos clases (puntos dentro y fuera de la hiper-esfera).

Este procedimiento funciona muy bien cuando las clases son esféricamente separables. En caso de tener tres o más clases el procedimiento es similar. En este caso, el conjunto de puntos \mathbf{p}^i $_{i=1}^l$ serán todos los patrones de la k -ésima clase, y el conjunto de puntos \mathbf{p}^j $_{j=l+1}^m$ serán todos los patrones de las demás clases. La k -ésima hiper-esfera S^k se encuentra al resolver la expresión (1.21). El mismo procedimiento se debe realizar para las otras clases.

4.1.3 Fase de aprendizaje

La fase de aprendizaje de una memoria asociativa consiste en almacenar las asociaciones entre los patrones de entrada y sus correspondientes patrones de salida. En el caso de las memorias asociativas geométricas, la fase de aprendizaje consiste en crear las vecindades esféricas para cada clase.

Una memoria asociativa geométrica es, por lo tanto, una matriz \mathbf{M} de tamaño $q \times n + 2$, donde q es el número total de clases. El k -ésimo renglón de \mathbf{M} es la k -ésima hiper-esfera. \mathbf{M} se puede observar en la Figura 2, donde C^k y γ^k son, respectivamente, el centro y el radio de la k -ésima hiper-esfera.

Las hiper-esferas se pueden crear usando alguno de los procedimientos mencionados en la sección 4.1, aunque se recomienda utilizar el procedimiento descrito en la sección 4.1.1.

4.1.4 Fase de clasificación

Por otro lado, para la fase de clasificación de una memoria asociativa geométrica se realizará usando la idea principal, presentada originalmente en (Banarer, Perwass y Sommer 2003), pero con ciertos cambios para considerar el caso de patrones que se encuentren dentro de dos hiper-esferas.

$$\mathbf{M} = \begin{bmatrix} S^1 \\ S^2 \\ \vdots \\ S^q \end{bmatrix} = \begin{bmatrix} S^1 = C^1 - \frac{1}{2}(\gamma^1)^2 e_\infty \\ S^2 = C^2 - \frac{1}{2}(\gamma^2)^2 e_\infty \\ \vdots \\ S^q = C^q - \frac{1}{2}(\gamma^q)^2 e_\infty \end{bmatrix}$$

Figura 2. Memoria asociativa geométrica después de la fase de aprendizaje.

Primero se debe aplicar un producto interior entre un vector (fundamental o de prueba) x y la propia memoria asociativa \mathbf{M} para obtener un vector u de tamaño m . Dicho vector contiene todos los productos interiores entre el patrón x y todas las hiper-esferas. La k -ésima componente de u está dada por:

$$u_k = \mathbf{M}_k \cdot X = S^k \cdot X. \quad (1.22)$$

En este caso X es la representación conforme del patrón x . Si x está dentro de una hiper-esfera, la ecuación (1.22) regresa un valor positivo o cero en la posición de esa hiper-esfera y un valor negativo cuando se encuentra fuera de la hiper-esfera. En algunos casos (principalmente cuando se trata de patrones afectados con ruido) x podría estar dentro de dos o más hiper-esferas o podría estar fuera de todas las hiper-esferas. Para decidir a cuál hiper-esfera pertenece un patrón dado, se debe utilizar el siguiente re-mapeo:

$$v_k = \begin{cases} -\infty; & \text{si } u_k < 0 \\ u_k - \gamma^k{}^2; & \text{otro caso} \end{cases}. \quad (1.23)$$

Este re-mapeo se debe realizar para $k = 1, \dots, q$. Finalmente, el identificador de clase se obtiene aplicando una función de argumento máximo:

$$j = \arg \max_k v_k \quad k=1, \dots, q. \quad (1.24)$$

Como se puede observar, cuando X está afuera de la k -ésima hiper-esfera, la expresión (1.23) regresa $-\infty$ y cuando X está dentro de la k -ésima hiper-esfera, regresa la distancia (con signo menos) entre el vector X y el centro de la k -ésima hiper-esfera C^k . Por tanto, el patrón x será clasificado por una hiper-esfera cubriendo su representación conforme y con ayuda de la expresión (1.24) éste será clasificado por la hiper-esfera con centro más cercano a x .

En algunos casos $v_k = -\infty \forall_{k=1, \dots, q}$, esto es, X está afuera de todas las hiper-esferas. Cuando se aplique la expresión (1.24), ésta no puede regresar algún valor. En este punto, se pueden tener dos situaciones. La primera es que x no pertenece a ninguna clase. La segunda opción es usar la expresión (1.24) directamente sobre $u_k - \gamma^k$, aquí la memoria se comportaría como un clasificador de distancia mínima, pero no se tomaría en cuenta el uso de las vecindades esféricas.

Hay que apuntar que la fase de clasificación no depende de la fase de aprendizaje.

4.1.5 Condición para clasificación perfecta

Por la **¡Error! No se encuentra el origen de la referencia.**, cuando una memoria asociativa clasifica correctamente el conjunto fundamental, entonces la memoria presenta clasificación perfecta. Sea \mathbf{M} una memoria asociativa geométrica entrenada con el método descrito en la página 3.

Teorema 2. *Sean q clases esféricamente separables en \mathbb{R}^n y sea \mathbf{M} una memoria asociativa geométrica entrenada para esas clases, entonces \mathbf{M} presenta recuperación perfecta.*

Demostración. *Sea k un índice de clase cuya hiper-esfera S^k es la k -ésima componente de \mathbf{M} y sea \mathbf{p} un patrón fundamental de la clase k , y sea j un índice $j=1, \dots, q$ tal que $j \neq k$. La hiper-esfera S^k se obtuvo por medio de la expresión (1.21). Entonces, de acuerdo a las condiciones (1.4) y (1.5), $\mathbf{p} \cdot S^k \geq 0$ porque \mathbf{p} es un patrón de clase k , y $\mathbf{p} \cdot S^j < 0$. Cuando se aplica la expresión (1.23), el vector \mathbf{u} tiene un número positivo o cero en la posición k y $-\infty$ en las otras posiciones. Por lo tanto la expresión (1.24) regresa k . Lo mismo sucede con todos los demás patrones en todas las demás clases ■.*

4.1.6 Condición para clasificación robusta

De acuerdo a la ~~¡Error! No se encuentra el origen de la referencia.~~, una memoria asociativa presenta clasificación robusta cuando ésta clasifica patrones afectados con ruido.

La robustez de una memoria asociativa geométrica depende en gran medida del tamaño de su radio; la memoria puede clasificar cualquier patrón como perteneciente a su clase cuando el patrón se encuentre dentro de la hiper-esfera. Los patrones localizados fuera de una hiper-esfera específica no serán clasificados como pertenecientes a dicha hiper-esfera.

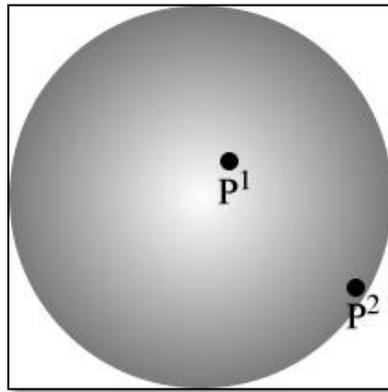


Figura 3. Ejemplo de dos patrones en \mathbb{R}^3 dentro de una hiper-esfera, el patrón P^2 , más cercano al borde, acepta menos cantidad de ruido que P^1 , que está más cercano al centro.

La cantidad de ruido que puede admitir un patrón fundamental depende de la posición del mismo con respecto al centro y al borde de la hiper-esfera. Los patrones que se encuentran más cerca del centro pueden admitir más cantidad de ruido que aquellos localizados más cerca del borde (Figura 3).

Teorema 3. Sea \tilde{p} un patrón fundamental afectado con ruido, tal que $\tilde{p} = p + r$, donde r es el ruido y sea M una memoria asociativa geométrica donde la hiper-esfera S^i es la hiper-esfera de la clase a la cual pertenece p y m es el número de clases. Si se cumple la expresión (4.25) para algún $i = 1, \dots, m$, entonces \tilde{p} será clasificado por M a la clase más cercana a él.

$$S^i \cdot \tilde{P} \geq -\frac{\gamma^i{}^2}{2}. \quad (1.25)$$

Demostración. *Mediante la expresión ¡Error! No se encuentra el origen de la referencia., se obtiene la expresión (4.26), al obtener el producto interior entre una la esfera y el patrón afectado con ruido.*

$$S^i \cdot \tilde{P} = \frac{1}{2} \gamma^i{}^2 - \mathbf{c}^i - \tilde{\mathbf{p}}{}^2 \quad (1.26)$$

Mediante esta expresión y la expresión (1.25) se obtiene el siguiente desarrollo:

$$\begin{aligned} -\frac{1}{2} \mathbf{c}^i - \tilde{\mathbf{p}}{}^2 &\geq -\frac{\gamma^i{}^2}{2} \\ -\mathbf{c}^i - \tilde{\mathbf{p}}{}^2 &\geq -\gamma^i{}^2 \\ \mathbf{c}^i - \tilde{\mathbf{p}}{}^2 &\leq \gamma^i{}^2 \\ 0 &\leq \gamma^i{}^2 - \mathbf{c}^i - \tilde{\mathbf{p}}{}^2 \\ \gamma^i{}^2 - \mathbf{c}^i - \tilde{\mathbf{p}}{}^2 &\geq 0 \\ 2 S^i \cdot \tilde{P} &\geq 0 \\ S^i \cdot \tilde{P} &\geq 0 \end{aligned} \quad (1.27)$$

Por lo tanto, $\tilde{\mathbf{p}}$ estará dentro de alguna esfera S^i y mediante la expresión (1.24), se garantiza que dicho patrón se encuentre dentro de esfera con centro más cercano ■.

Note que el Teorema 3 no depende del tipo de ruido que afecta al patrón.

4.2 Modo para restauración de patrones

A continuación se presenta un nuevo modo de operación de las memorias asociativas geométricas, basado en las ideas mostradas en la sección 4.1. En este caso se pretende resolver el problema de la restauración de patrones. La idea general es asignar a cada patrón del conjunto fundamental una vecindad esférica, esa funciona como una zona

de encendido y se pueden restaurar patrones dentro de esa vecindad. El patrón recuperado será, precisamente, el centro de la hiper-esfera.

La principal ventaja de este modelo es su capacidad de restaurar, de forma directa, patrones afectados con ruido mezclado. Además, la cantidad de ruido que un patrón puede admitir se encuentra dada por el radio de la vecindad esférica.

A continuación se presenta una nueva forma de construir las memorias asociativas geométricas para que puedan restaurar patrones afectados con ruido.

4.2.1 Creación de las vecindades esféricas

Al igual que en el modo de clasificación, la construcción de las vecindades esféricas puede hacerse, de forma sencilla, tomando el patrón mismo como centro de la hiper-esfera y el radio se calcula por medio de la distancia entre este patrón y los demás. De esta forma se garantiza que no habrá solapamiento entre las hiper-esferas. Sin embargo, al igual que con el modo de clasificación, este procedimiento es computacionalmente costoso, así que de la misma que en la sección 4.1.1 se utilizó una aproximación por mínimos cuadrados para encontrar una hiper-esfera óptima.

Sea $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^m$ un conjunto fundamental de patrones en \mathbb{R}^n , el problema es encontrar una hiper-esfera óptima S con el mínimo error cuadrático, tal que \mathbf{p}^k se encuentra dentro de S y $\mathbf{p}^j \mid j=1, \dots, m, j \neq k$ se encuentran fuera de ella. La solución se puede encontrar aplicando, de nuevo, la expresión (1.3) con sus respectivas restricciones. Podría pensarse que este problema podría resolverse de la misma manera que el modo de clasificación, mediante la expresión (1.21); sin embargo si se utiliza esta expresión como tal, el centro de la hiper-esfera podría no ser el patrón \mathbf{p}^k . La idea, como se mencionó anteriormente, es regresar, como patrón restaurado, el centro de la hiper-esfera y con este procedimiento, esto no sería posible.

La solución es agregar una nueva restricción a la expresión (1.3), en la cual el centro de la hiper-esfera es igual a \mathbf{p}^k :

$$P^k = C \Leftrightarrow \mathbf{p}^k = \mathbf{c}, \quad (1.28)$$

donde C es el centro de la hiper-esfera y P^k es el patrón a entrenar, ambos en su representación conforme. Entonces, partiendo de la expresión (1.7) (omitiendo el tercer término de la expresión sin pérdida de generalidad) y tomando en cuenta la restricción (1.28) se obtiene:

$$\begin{aligned}
\sum_{i=1}^m P^i \cdot S^2 &= \sum_{i=1}^m p^i \cdot p^k - S_{n+1}^2 + \sum_{i=1}^m -p^i \cdot p^k + S_{n+1} p^{i^2} \\
&= \sum_{i=1}^m p^i \cdot p^k^2 - 2 p^i \cdot p^k S_{n+1} + S_{n+1}^2 \\
&\quad - \sum_{i=1}^m p^i \cdot p^k - S_{n+1} p^{i^2} \\
&= \sum_{i=1}^m p^i \cdot p^k^2 - \sum_{i=1}^m 2 p^i \cdot p^k S_{n+1} + \sum_{i=1}^m S_{n+1}^2 \\
&\quad - \sum_{i=1}^m p^i \cdot p^k p^{i^2} + \sum_{i=1}^m S_{n+1} p^{i^2}.
\end{aligned} \tag{1.29}$$

En este caso el primer y el cuarto término de esta expresión se pueden omitir, debido a que no dependen de S . De esta forma se obtiene el siguiente desarrollo:

$$\begin{aligned}
\sum_{i=1}^m P^i \cdot S^2 &= -\sum_{i=1}^m 2 p^i \cdot p^k S_{n+1} + \sum_{i=1}^m S_{n+1}^2 \\
&\quad + \sum_{i=1}^m S_{n+1} p^{i^2}, \\
&= -2S_{n+1} \sum_{i=1}^m p^i \cdot p^k + m S_{n+1}^2 + S_{n+1} \sum_{i=1}^m p^{i^2}, \\
&= S_{n+1} \left(2 \sum_{i=1}^m p^i \cdot p^k + \sum_{i=1}^m p^{i^2} \right) + m S_{n+1}^2, \\
&= m S_{n+1}^2 + S_{n+1} \sum_{i=1}^m 2 p^i \cdot p^k + p^{i^2}.
\end{aligned} \tag{1.30}$$

Finalmente, sean $\mathbf{y} = \left(\sum_{i=1}^m 2 \mathbf{p}^i \cdot \mathbf{p}^k + \mathbf{p}^i \cdot \mathbf{p}^i \right)$ y $x = S_{n+1}$. Al igual que en la sección 4.1.1, se tiene una expresión en notación matricial euclidiana como en la siguiente expresión:

$$\min_x \mathbf{x} \mathbf{m} \mathbf{x} + \mathbf{y} \mathbf{x} . \quad (1.31)$$

A diferencia del modelo de clasificación, en este caso la variable a optimizar es, simplemente S_{n+1} . La condición para puntos dentro de la hiper-esfera es:

$$\begin{aligned} P^i \cdot S &\geq 0 \\ \mathbf{p}^i \cdot \mathbf{c} - S_{n+1} - \frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i &\geq 0 \\ \mathbf{p}^i \cdot \mathbf{p}^i - S_{n+1} - \frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i &\geq 0 \\ -S_{n+1} + \frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i &\geq 0 \\ -S_{n+1} &\geq -\frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i \\ S_{n+1} &\leq \frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i . \end{aligned} \quad (1.32)$$

Por su parte, la condición para los puntos fuera de la hiper-esfera es:

$$\begin{aligned} P^j \cdot S &< 0 \\ \mathbf{p}^j \cdot \mathbf{c} - S_{n+1} - \frac{1}{2} \mathbf{p}^j \cdot \mathbf{p}^j &< 0 \\ \mathbf{p}^j \cdot \mathbf{p}^i - S_{n+1} &< \frac{1}{2} \mathbf{p}^i \cdot \mathbf{p}^i \\ -S_{n+1} &< \frac{1}{2} \mathbf{p}^j \cdot \mathbf{p}^j - \mathbf{p}^j \cdot \mathbf{p}^i \\ -S_{n+1} &\leq \frac{1}{2} \mathbf{p}^j \cdot \mathbf{p}^j - \mathbf{p}^j \cdot \mathbf{p}^i + \varepsilon . \end{aligned} \quad (1.33)$$

De la misma forma que el modelo de clasificación, las restricciones (1.32) y (1.33) se pueden convertir a una notación matricial euclidiana, como en la expresión (1.20) y

obtener, finalmente, una expresión similar a (1.21); sin embargo, al ser solamente una variable (S_{n+1}) a optimizar, ésta se puede obtener directamente al calcular la primera derivada de (1.31), pero tomando en cuenta el intervalo entre las dos restricciones.

La hiper-esfera óptima S se encuentra al resolver la expresión (1.21), solo hay que despejar el valor del radio (γ) de S_{n+1} . El centro de la hiper-esfera será el propio punto p^i . En este caso, la solución S es una superficie de separación que permite separar un punto en particular, de todos los demás.

4.2.2 Fase de aprendizaje

Como en el modo de clasificación, la fase de aprendizaje de una memoria asociativa consiste en almacenar las asociaciones entre los patrones de entrada y sus correspondientes patrones de salida. En este caso, la fase de aprendizaje consiste en crear las vecindades esféricas para cada patrón.

Al igual que en el modo de clasificación, la memoria asociativa geométrica es una matriz \mathbf{M} de tamaño $q \times n+2$, donde q es el número de patrones. El k -ésimo renglón de \mathbf{M} es la k -ésima hiper-esfera. \mathbf{M} se puede observar en la Figura 2.

Las hiper-esferas se pueden construir al usar el procedimiento descrito en la sección 4.1.2. Este mismo procedimiento debe realizarse para todos los patrones en el conjunto fundamental.

4.2.3 Fase de restauración

La restauración de un patrón se puede realizar con los mismos pasos que en la Fase de clasificación del modo de clasificación, mostrados en la sección 4.1.4, excepto el último paso, mientras que en el modelo de clasificación se regresa un índice, en el modelo de restauración se regresa el centro de la j -ésima hiper-esfera, como el patrón restaurado, donde j es el valor obtenido por medio de la expresión (1.24).

Al igual que en el modelo de clasificación, un patrón afectado con ruido será restaurado por la hiper-esfera que cubra su representación conforme y con centro más cercano a dicho patrón. En este modelo, las hiper-esferas funcionan como “atractores” del patrón a restaurar.

De la misma forma que en el modelo de clasificación, la fase de restauración de este modo no depende de la fase de aprendizaje.

4.2.4 Condición para restauración perfecta

La condición para restauración perfecta (**¡Error! No se encuentra el origen de la referencia.**) de una memoria asociativa geométrica se enuncia en el Teorema 4. Sea

\mathbf{p}^i un conjunto fundamental de patrones en \mathbb{R}^n y sea \mathbf{M} su memoria asociativa geométrica construida como se mostró en las secciones 4.2.1 y 4.2.2.

Teorema 4. *\mathbf{M} presenta restauración perfecta.*

Demostración. *Sea S^i la hiper-esfera correspondiente al i -ésimo patrón (\mathbf{p}^i). S^i se obtuvo usando la expresión (1.3) y siguiendo el procedimiento de la sección 4.2.1, por lo tanto se cumplen las condiciones (1.4) y (1.5). Al aplicar la expresión (1.23) el vector \mathbf{u} tiene un número positivo en la posición i y $-\infty$ en las demás posiciones. Entonces, si se aplica la expresión (1.24), ésta regresa i . Finalmente el patrón restaurado es, por lo tanto es el centro de la hiper-esfera S^i , es decir \mathbf{p}^i ■.*

4.2.5 Condición para restauración robusta

Al igual que en el modelo anterior, la robustez (**¡Error! No se encuentra el origen de la referencia.**) de una memoria asociativa geométrica depende del tamaño de su radio. La cantidad de ruido que esta memoria puede admitir se enuncia en el Teorema

5. Sea \mathbf{p}^i un conjunto fundamental de patrones en \mathbb{R}^n y sea \mathbf{M} su memoria asociativa geométrica construida como se mostró en las secciones 4.2.1 y 4.2.2. Sea, además, $\tilde{\mathbf{p}}^i = \mathbf{p}^i + \mathbf{r}$ un patrón fundamental afectado con ruido.

Teorema 5. *Si se cumple la expresión (4.34), donde γ^i es el radio de la i -ésima hiper-esfera. El patrón $\tilde{\mathbf{p}}^i$ puede ser restaurado por \mathbf{M} .*

$$r^2 \leq \gamma^i{}^2 \tag{1.34}$$

Demostración. Sea S^i la hiper-esfera del patrón i -ésimo, obtenida mediante el método descrito en la sección 4.2.1. El ruido del patrón $\tilde{\mathbf{p}}^i$ se puede expresar de la siguiente manera:

$$\begin{aligned}\tilde{\mathbf{p}}^i &= \mathbf{p}^i + \mathbf{r} \\ \mathbf{r} &= \tilde{\mathbf{p}}^i - \mathbf{p}^i \\ -\mathbf{r} &= \mathbf{p}^i - \tilde{\mathbf{p}}^i \\ r^2 &= \|\mathbf{p}^i - \tilde{\mathbf{p}}^i\|^2.\end{aligned}\tag{1.35}$$

Si se aplica el producto interior entre la hiper-esfera S^i y la representación conforme de $\tilde{\mathbf{p}}^i$, por la ecuación **¡Error! No se encuentra el origen de la referencia.**, se obtienen las expresiones (4.36).

$$\begin{aligned}2 \tilde{\mathbf{P}} \cdot S^i &= \gamma^2 - \|\mathbf{c}^i - \tilde{\mathbf{p}}^i\|^2 \\ &= \gamma^2 - \|\mathbf{p}^i - \tilde{\mathbf{p}}^i\|^2 \\ &= \gamma^2 - r^2.\end{aligned}\tag{1.36}$$

Entonces:

- Si $r^2 > \gamma^2$ entonces $\tilde{\mathbf{P}}^i \cdot S^i < 0$.
- Si $r^2 \leq \gamma^2$ entonces $\tilde{\mathbf{P}}^i \cdot S^i \geq 0$.

Por lo tanto, el patrón $\tilde{\mathbf{p}}^i$ será restaurado por \mathbf{M} si y sólo si $r^2 \leq \gamma^2$ ■.

Hay que recalcar que el Teorema 5 no depende del tipo de ruido utilizado.

4.3 Modo para clasificación no supervisada

Originalmente las memorias asociativas geométricas para clasificación de patrones fueron desarrolladas para funcionar de forma supervisada. Sin embargo, en esta sección se mostrará un algoritmo de clasificación no supervisada para estas memorias.

Se presentará una variante del famoso algoritmo *k-medias* (del inglés *k-means*) (McQueen 1967), pero usando operadores del álgebra geométrica conforme para construir los *clusters* (agrupamientos) respectivos. Además, al igual que en los modos

anteriores de operación, se puede utilizar el producto interior para saber si un patrón dado se encuentra dentro o fuera de alguno de los agrupamientos. También con el producto interior, es muy fácil conocer la distancia entre dos puntos.

4.3.1 Clasificación no supervisada con álgebra geométrica

El algoritmo propuesto se basa en las ideas mencionadas en las líneas anteriores, pero se usa en el marco de trabajo del álgebra geométrica conforme. A diferencia del algoritmo k -medias, los agrupamientos se encontrarán separando los puntos asociados a un centro específico de los otros puntos al resolver el problema de optimización mostrado en la sección 4.1.1; es decir, al usar una memoria asociativa geométrica. La principal ventaja de este procedimiento es que los nuevos centros se calculan automáticamente.

El método de la sección 4.1.1 toma en cuenta los puntos dentro y fuera de una hiper-esfera de clasificación para un mejor rendimiento. Al igual que en el modo de restauración, las hiper-esferas en la clasificación no supervisada funcionan como atractores para los puntos que están dentro y como contractores para los puntos que están afuera.

Otra diferencia es el cálculo de la distancia entre los centros y los patrones correspondientes. Este procedimiento se realiza con la memoria asociativa geométrica, formada por los centros correspondientes, y con un producto interior.

Durante la primera iteración del algoritmo, los centros se generan de manera aleatoria, como en el algoritmo de k -medias; en las siguientes iteraciones los centros se calculan automáticamente al construir las hiper-esferas correspondientes.

4.3.2 Nuevo algoritmo de agrupamiento

Básicamente el algoritmo es como sigue. Dado un conjunto de puntos \mathbf{p}^i en \mathbb{R}^n , y sea k el número de agrupaciones a generar para esos puntos:

1. Cambiar cada punto \mathbf{p}^i su respectiva representación conforme P^i por medio de la expresión **¡Error! No se encuentra el origen de la referencia..**

2. Establecer k puntos aleatorios, en su representación conforme: C^j donde $j = 1, \dots, k$.
3. Construir la memoria $M = [C^1, \dots, C^k]^t$.
4. Generar k agrupaciones asociando cada punto p^i con el centro más cercano. Este procedimiento se puede realizar aplicando:

$$\arg \max \mathbf{M} \cdot P^i \quad (1.37)$$

5. Encontrar k superficies esféricas S^k por cada agrupación, por medio de la expresión (1.3) y sus respectivas restricciones.
6. Los nuevos valores de los k centros serán, precisamente, los centros de cada hiper-esfera.
7. Repetir desde el paso 3 hasta que los centros no cambien o se llegue a un número determinado de iteraciones.

Como se puede observar, en este algoritmo el cálculo de las agrupaciones se realiza por medio de un simple producto interior y una función de máximo; a diferencia del método tradicional que, para encontrar el centro más cercano utiliza un algoritmo del vecino más cercano.

Con este algoritmo se puede presentar el caso, en algunas situaciones, de que algunos puntos se encuentren fuera de todas las hiper-esferas. Esto, puede ser una ventaja como una desventaja, dependiendo de las características del problema. En algunas ocasiones, estos puntos podrían ser considerados como ruido y se pueden eliminar sin perder información esencial sobre la naturaleza del problema. Sin embargo, para algunas otras situaciones de agrupación estos puntos pueden ser esenciales para resolver el problema.