



Mixtura de Gaussianas

Aprendizaje Automático

Resultados obtenidos

En esta primera parte del proyecto hemos estudiado el clasificador de mixtura de gaussianas. Para realizar la estimación de los parámetros de una mixtura de gaussianas en el marco del algoritmo EM, lo primero que hemos hecho ha sido implementar el paso M.

En el paso M hemos estimado los parámetros del modelo teniendo una estimación del grado de pertenencia de cada muestra de entrenamiento a cada componente. Los pasos que hemos seguido vienen dados por las siguientes fórmulas:

$$\hat{p}_{ck}^{(t+1)} = \frac{1}{N_c} \sum_{n: c_n=c} z_{nk}^{(t)} \quad (7)$$

$$\hat{\mu}_{ck}^{(t+1)} = \frac{1}{\sum_{n: c_n=c} z_{nk}^{(t)}} \sum_{n: c_n=c} z_{nk}^{(t)} \mathbf{x}_n \quad (8)$$

$$\hat{\Sigma}_{ck}^{(t+1)} = \frac{1}{\sum_{n: c_n=c} z_{nk}^{(t)}} \sum_{n: c_n=c} z_{nk}^{(t)} (\mathbf{x}_n - \hat{\mu}_{ck})(\mathbf{x}_n - \hat{\mu}_{ck})^t \quad (9)$$

La implementación de estos pasos en el fichero `mixgaussian.m` corresponde a las siguientes líneas de código:

```
sumz=sum(z);  
  
% Equac 7  
pkGc{ic}=sumz/Nc;  
  
% Equac 8  
mu{ic}=(Xc'*z)./sumz;  
  
% Equac 9  
for k=1:K  
    mc = ((Xc-mu{ic}(:,k))')' * ((Xc-mu{ic}(:,k))'*z(:,k)))/sumz(k);  
    sigma(ic,k) = alpha * mc + (1-alpha)*eye(D);  
end
```

Una vez implementado esto hemos hecho una primera comprobación comparando su funcionamiento para el caso de una única componente por mixtura con el clasificador gaussiano, lo cual se encuentra implementado en el fichero `mixgaussian-exp.m`

Los resultados obtenidos han sido los siguientes:

alpha	Ks	dv-err			
It		oL	L	errX	errY
1	-Inf	-620842.17812	3.92	4.27	
2	-620842.17812	-313498.91244	3.92	4.27	
3	-313498.91244	-313498.91244	3.92	4.27	
1.0e-04	1	4.267			

Con esta comprobación terminada hemos pasado a la evaluación del error de clasificación en función del número de componentes por mixtura del clasificador ($K = 1, 2, 5, 10, 20, 50, 100$), para un número de dimensiones PCA variable ($D = 1, 2, 5, 10, 20, 50, 100$) y hemos probado algunos de los valores de α de suavizado *flat smoothing* que mejores resultados han proporcionado al clasificador gaussiano, es decir, $\alpha = 1e-3, 1e-4$ y $1e-5$.

Hemos elegido estos valores de α debido al alto coste computacional del clasificador. Es por eso por lo que hemos cogido el valor con mejores resultados ($\alpha = 1e-4$) y los dos valores más próximos a él ($\alpha = 1e-3$ y $\alpha = 1e-5$).

Los datos obtenidos al ejecutar el script `pca+mixgaussian-exp.m` con los valores de α mencionados anteriormente son los siguientes:

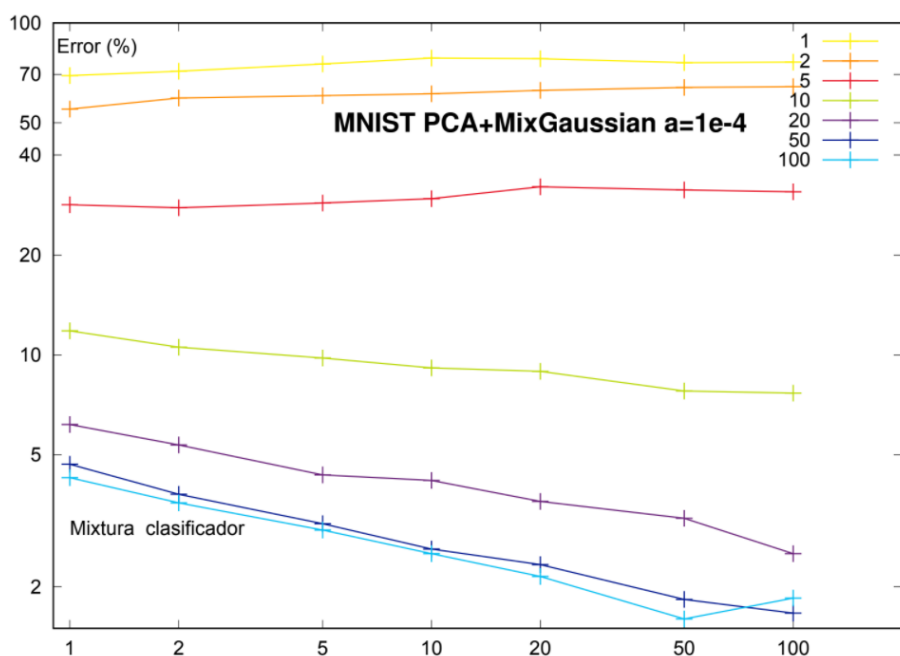


Figure 1. Resultados de `pca+mixgaussian-exp.m` con $\alpha = 1e-4$

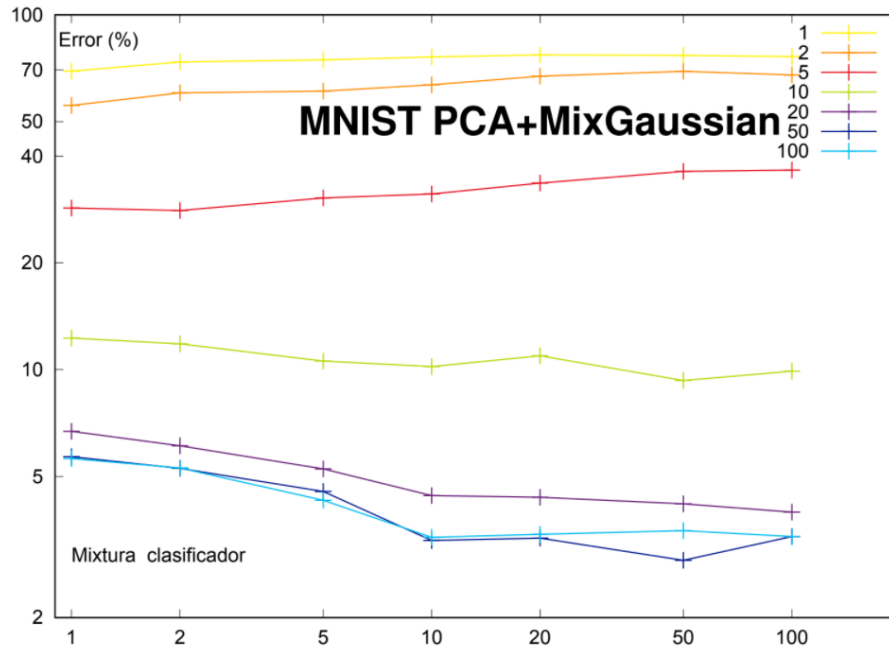


Figure 2. Resultados de `pca+mixgaussian-exp.m` con $\alpha = 1e-3$

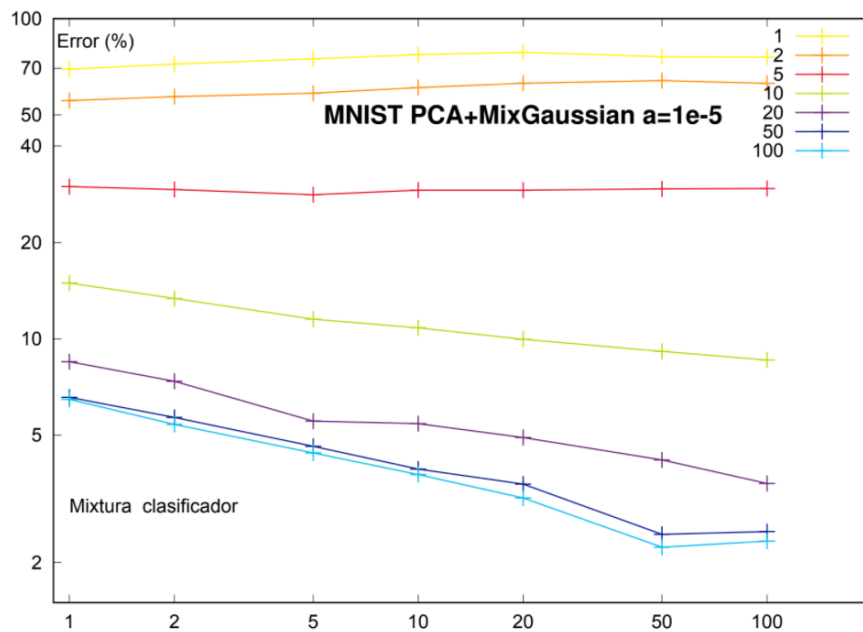


Figure 3. Resultados de `pca+mixgaussian-exp.m` con $\alpha = 1e-5$

En estas gráficas observamos la relación entre los distintos valores de α y el error que generan. De esta forma podemos confirmar que el valor de α $1e-4$ es el que nos produce el menor error, por lo que será el valor que utilizemos para ejecutar `pca+mixgaussian-eva.m`.

A partir de la ejecución del programa `pca+mixgaussian-exp.m` hemos extraído los parámetros para las mejores tasas de error.

En concreto, para $pca=50$ y $Ks=50$ tenemos un error de 1.870%

Mientras que para $pca=100$ y $Ks=50$ el error es de 1.920%

Además, hemos obtenido los siguientes intervalos de confianza:

Para 1.870 -> [1.604,2.136]

Para 1.920 -> [1.651,2.189]

A modo de prueba, hemos ejecutado también el programa `pca+mixgaussian-eva.m` con los parámetros $pca=50$, $Ks=20$ y $pca=50$ y $Ks=100$, obteniendo un error de 1.900% y 1.620%, respectivamente.

Para comparar los resultados del clasificador de mixtura gaussiana con el gaussiano, nos hemos fijado en el valor del error correspondiente en la página web de MNIST de “40 PCA + quadratic classifier”. Ahí, podemos observar que hemos obtenido un valor inferior del error para todos los valores que hemos calculado. En concreto, el valor proporcionado por MNIST es 3.3%.

También podemos comparar con los resultados con los que calculamos el año pasado en la asignatura PER, ya que obtuvimos también en `pca+gaussian-eva.m` que el mínimo error era del 3.7%. Por tanto, el error del clasificador mixtura de gaussianas es menor que el del clasificador gaussiano.

Según los datos de la página de MNIST, este clasificador también sería mejor que otros clasificadores vistos en la asignatura PER, como por ejemplo los clasificadores lineales. El error reportado para “linear classifier (1-layer NN)” es del 12%, muy superior al que hemos obtenido en nuestro clasificador.