

Factorizaciones matriciales aproximadas aplicadas en el reconocimiento de género

Factorización no negativa de Matrices y una versión local

Jorge Luis Ramos Zavaleta

Centro de Investigación en Matemáticas (CIMAT A.C.), Unidad Monterrey

jorge.ramos@cimat.mx



Resumen

Un rostro puede ser representado conceptualmente como una conjunto de partes distribuidas de manera esparcida: ojos, nariz, boca, etc. Los humanos tenemos la habilidad de encontrar patrones en las caras lo que nos permite diferenciar entre el rostro de un hombre y el de una mujer. El siguiente reporte presenta los resultados de aplicar NMF(non-negative matrix factorization), LNMF(Local NMF) y PCA a diversas imágenes de rostros de hombres y mujeres con el fin de generar un clasificador de género comparando los 3 métodos para conocer su efectividad en dicha tarea.

Introducción

Debido a los problemas que se generan en altas dimensiones ya sea por redundancias o superposiciones de características, el análisis de objetos en el espacio de los datos originales puede ser computacionalmente exhaustivo. Por lo que problemas que requieren el análisis de imágenes en particular de rostros generalmente el problema es convertido a uno con un subespacio de dimensión más baja para que sea más fácilmente tratable.

Asimismo una representación en un subespacio suficiente bueno puede exponer estructuras escondidas y características inherentes al tipo de objetos con el que se quiere trabajar. Incluso puede considerarse alguna bondad de dicho subespacio en el que trabajar con diferentes variaciones como de iluminación o traslaciones sea mucho mas sencillo que si se tratara con el espacio original.

Esto nos lleva a establecer que la elección de una apropiada representación de los datos en un subespacio de menor dimensión que el original puede volverse imperativo para el éxito y eficacia de un sistema de reconocimiento de objetos o rostros como es nuestro caso.

PCA, NMF y LNMF

PCA es uno de los métodos más usados en estadística y es por ello que es muy conocido y aplicado para hacer reducción de dimensión así como para intentar encontrar características en datos multivariados. El método consiste en encontrar direcciones sobre las cuales las proyecciones de los datos generen la máxima varianza.

El enfoque de detección de rostros usando PCA es muy conocido y recibe el nombre de eigenfaces. Es un esquema muy simple el que se usa para aplicar PCA a diversos rostros. Para ello se genera una matriz con cada rostro vectorizado y colocado como una columna de dicha matriz después se calcula la imagen media y se le resta a cada imagen de la matriz para centrar las imágenes y evitar problemas dados por traslaciones. A esta matriz centrada se le calcula su factorización SVD y se considera el número de componentes a usar dependiendo el criterio que se elija para esto.

Al igual que PCA tanto NMF como su versión local son esquemas de aproximación de una matriz considerando un rango más bajo que el de la matriz original. Aun cuando se sabe que PCA (SVD) tiene la mejor aproximación de rango bajo de una matriz la naturaleza inherente de no negatividad de algunas estructuras de datos permite entrever la utilidad de estos esquemas multivariados que consideran dicha propiedad.



Figura 1: Izquierda: Eigenface obtenida utilizando PCA. Derecha: Atributos obtenidos usando NMF.

NonNegative Matrix Factorization (NMF)

A finales de los 1970's ya se había planteado la idea de usar una factorización de matrices con restricción de no negatividad, sin embargo el problema de optimización que esto conlleva es de tipo Np-hard por lo que se abandono la idea, y fue hasta que en 1999 se publica un resultado donde se propone una forma de resolver dicho problema de optimización haciendo uso de una regla de actualización multiplicativa.

A partir de dicho resultado se ha comenzado una extensa investigación este tipo de aproximación de rango bajo y de como generar extensiones del método para considerar problemas más complejos.

El problema que se intenta resolver es

$$\min D(X|BH), \quad s. a. B \geq 0 \text{ y } H \geq 0$$

Con

$$D(X|BH) = \sum_{i,j} \left(X_{ij} \log \frac{X_{ij}}{BH_{ij}} - X_{ij} + BH_{ij} \right)$$

Y para ello se hace uso de una actualización multiplicativa

$$h_{pj} = h_{pj} \sum_{i=1}^n \left(\frac{b_{ip} a_{ij}}{\sum_{k=1}^r b_{ik} h_{kj}} \right), \quad b_{ip} = b_{ip} \sum_{j=1}^m \left(\frac{a_{ij} h_{pj}}{\sum_{k=1}^r b_{ik} h_{kj}} \right)$$

$$b_{ip} = \frac{b_{ip}}{\sum_{k=1}^n b_{kp}}$$

Haciendo uso de estas funciones de actualización multiplicativa se tiene garantizado que la divergencia de Kullback-Leibler sea no creciente con cada actualización.

A partir del trabajo de Lee se han encontrado otras formas de resolver el problema de optimización. Sin embargo, uno de los problemas asociados a estos nuevos descubrimientos es que la factorización no es única.

Local NMF

Local NMF es una extensión del algoritmo original de NMF en el que se busca que se haga una búsqueda local para encontrar características locales en lugar de las características globales que encuentran PCA y NMF. En nuestro problema a resolver, permitiría encontrar características específicas que pueden determinar lo que define un rostro masculino o un rostro femenino.

Dado que se trata de una extensión de NMF entonces se tienen condiciones de optimalidad mayores que los de NMF. Dado que

$$A \approx BH$$

Definimos:

$$U = B^T B \text{ y } V = H H^T$$

Entonces para lograr que se establezca una búsqueda local se requiere que se cumplan las siguientes restricciones

- Máxima esparcidad en H.- Esta se logra minimizando la diagonal de U
- Máxima expresividad en B.-Esta se logra maximizando la traza de V
- Ortogonalidad de las bases.- Esta se logra minimizando la suma de las entradas de la matriz U

por lo que la divergencia de Kullback-Leibler toma la siguiente forma

$$D(X|BH) = \sum_{i,j} \left(X_{ij} \log \frac{X_{ij}}{BH_{ij}} - X_{ij} + BH_{ij} \right) + \alpha \sum_{i,j} u_{ij} - \beta \sum_i v_{ii}$$

Por lo que el problema de minimización a resolver es similar al que se resuelve con NMF y tambien puede resolverse usando una regla multiplicativa.

Comparación de los métodos en la detección de género

Un problema básico que se encuentra al querer hacer comparaciones entre métodos, como es nuestro caso, es que se deben considerar igualdad de condiciones para que los métodos no se vean comprometidos por alguna consideración que les pueda afectar directamente en su funcionamiento y en este caso en su nivel de precisión al clasificar.

Para esto se eligió realizar la detección del rostro en la imagen haciendo uso de del método de Cascada de Haar.

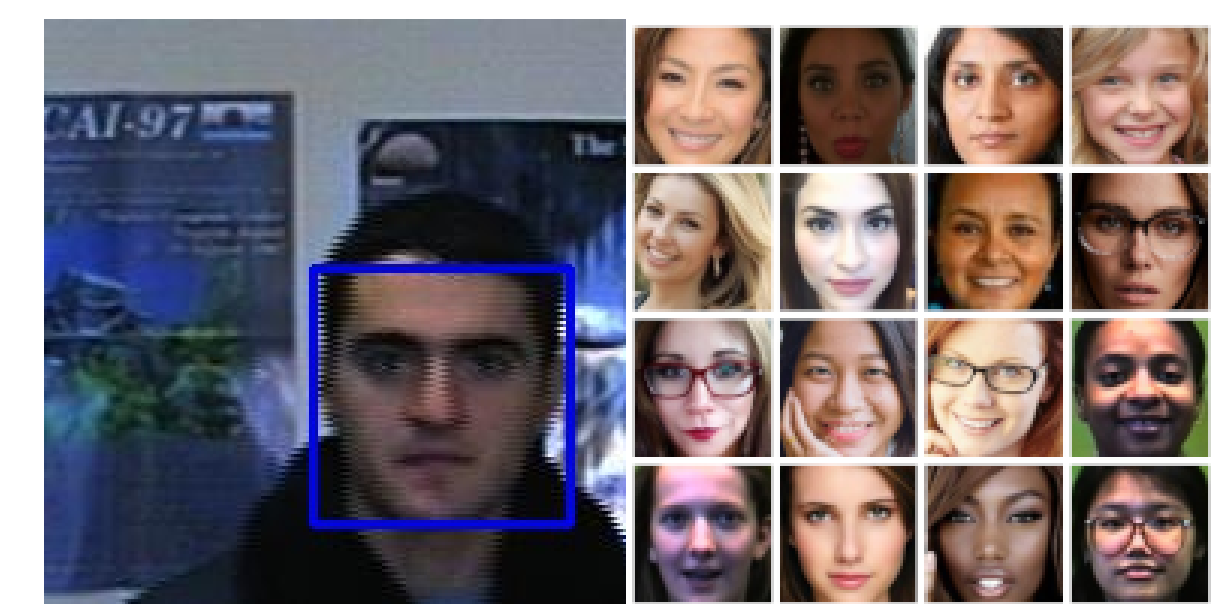


Figura 2: Izquierda: Rostro detectado haciendo uso de la cascada de Haar. Derecha: Muestra de rostros de mujeres obtenidos usando la Cascada de Haar.

Resultados

Para la implementación se consideraron 306 imágenes de hombres y mujeres en total. Para establecer la clasificación por género se genero un pequeño ensamble generando una base para las mujeres y otra para los hombres. A cada imagen nueva se le calcula la correlación contra cada elemento de la base de mujeres y se calcula la máxima correlación en esta base, de manera analoga se hace para la base de hombres, y el criterio de decisión es como sigue

$$\begin{aligned} \text{Max}(\text{Corr}(x, \text{Mujeres})) > \text{Max}(\text{Corr}(x, \text{hombres})) &\rightarrow M \\ \text{Max}(\text{Corr}(x, \text{Mujeres})) < \text{Max}(\text{Corr}(x, \text{hombres})) &\rightarrow H \end{aligned}$$

En caso de empate se elige aleatoriamente de manera uniforme.

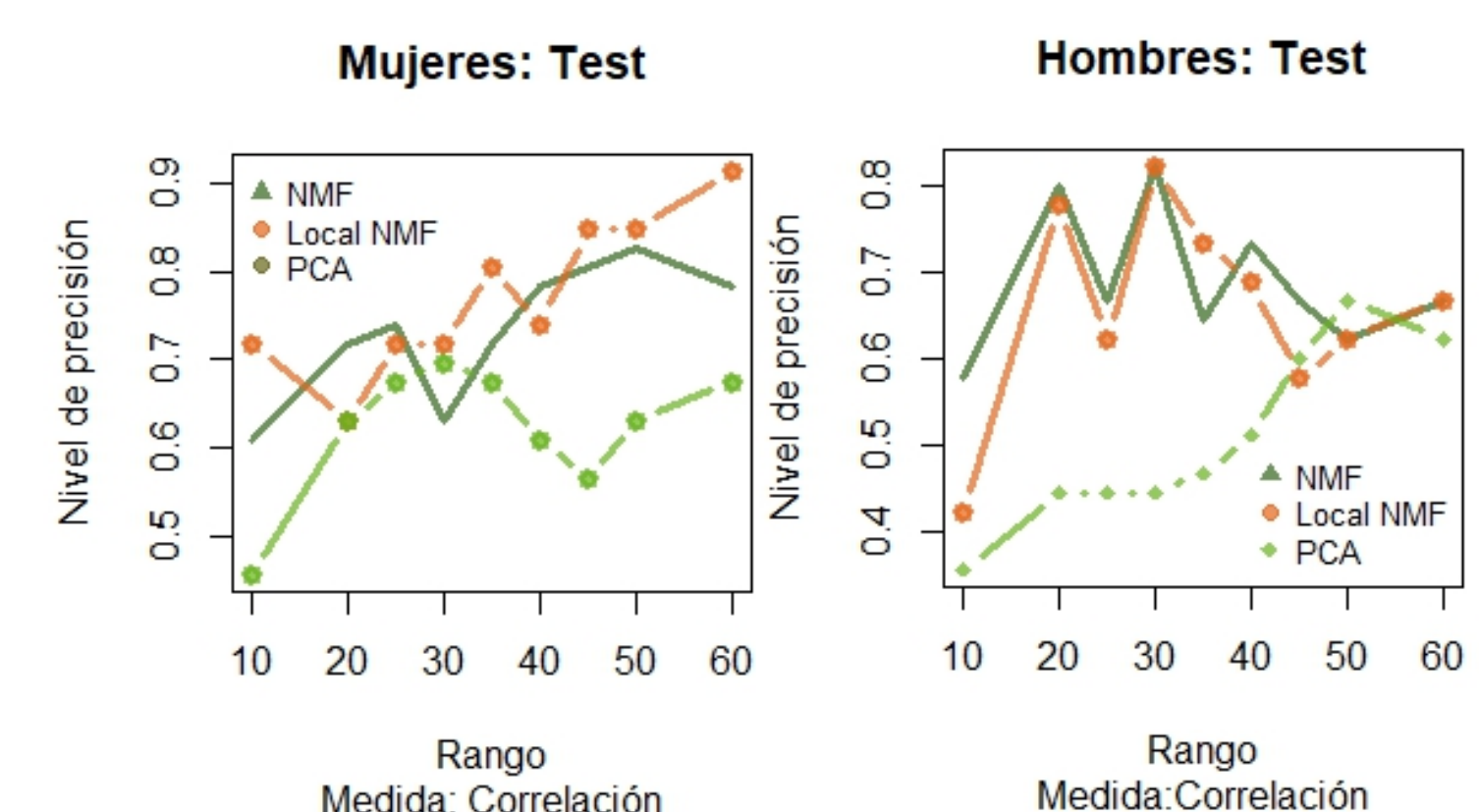


Figura 3: Izquierda: Nivel de precisión obtenido para el conjunto de rostros de mujer en el conjunto de prueba para diferentes rangos. Derecha: Nivel de precisión obtenido para el conjunto de prueba para hombres con distintos rangos.

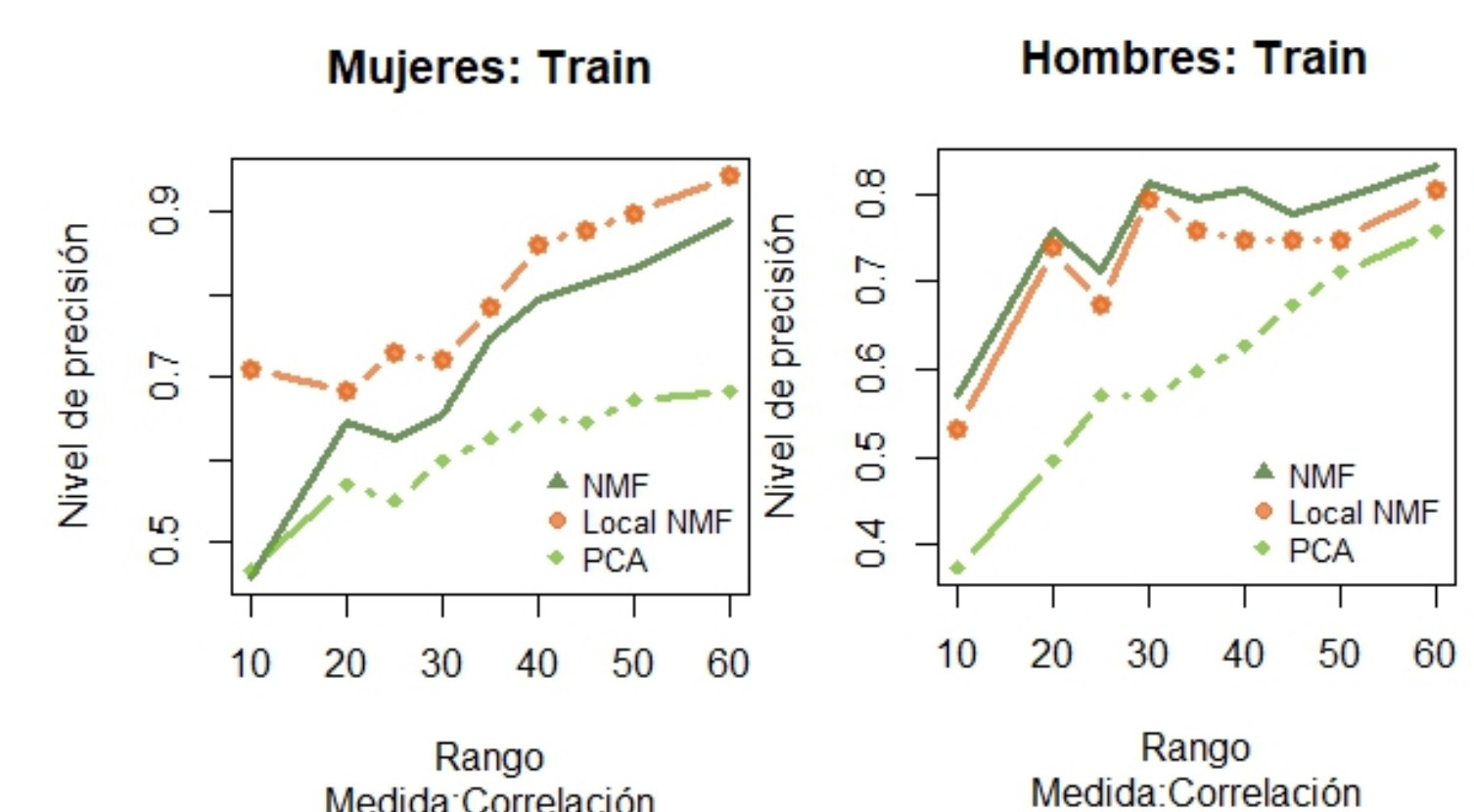


Figura 4: Izquierda: Nivel de precisión obtenido para el conjunto de rostros de mujer en el conjunto de entrenamiento para diferentes rangos. Derecha: Nivel de precisión obtenido para el conjunto de entrenamiento para hombres con distintos rangos.

Puede observarse que en el caso de mujeres LNMF logro encontrar atributos muy específicos en sus rostros por lo que en general supera a PCA y a NMF. Mientras que en el caso de los hombres parece no haber características tan distintivas en sus rostros por lo que clasifica al mismo nivel que NMF. Pero ambos casos logran superar en general a PCA mostrando la utilidad de este enfoque de no negatividad.

Referencias

- Ding, C. H., Li, T., & Jordan, M. I. (2010). Convex and semi-nonnegative matrix factorizations. IEEE transactions on pattern analysis and machine intelligence, 32(1), 45-55.
- Feng, T., Li, S. Z., Shum, H. Y., & Zhang, H. (2002). Local non-negative matrix factorization as a visual representation. In Development and Learning, 2002. Proceedings. The 2nd International Conference on (pp. 178-183). IEEE.
- Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755), 788.
- Li, S. Z., Hou, X. W., Zhang, H. J., & Cheng, Q. S. (2001). Learning spatially localized, parts-based representation. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. 1-1). IEEE.