Volume 10, Issue 6, June – 2025 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jun139

Clustering and Nonnegative Matrix Factorization: A Mathematical and Algorithmic Perspective

Dr. Mitat Uysal[1]

[1]Software Eng.Dept.-Dogus University

Publication Date: 2025/06/17

Abstract: Clustering is a fundamental task in machine learning and data analysis, enabling the discovery of inherent patterns within data. Nonnegative Matrix Factorization (NMF) has emerged as a powerful tool for clustering due to its ability to learn parts-based, interpretable representations. This article explores the theoretical foundations of clustering and NMF, their synergy, algorithmic formulations, and practical implementations. Experimental validation on synthetic data demonstrates the effectiveness of NMF-based clustering without using libraries such as sklearn or tensorflow.

Keywords: *NMF,Clustering,Machine Learning,Objective Function,Frobenius Norm.*

How to cite: Dr. Mitat Uysal; (2025) Clustering and Nonnegative Matrix Factorization: A Mathematical and Algorithmic Perspective. *International Journal of Innovative Science and Research Technology,* 10(6), 822-824. https://doi.org/10.38124/ijisrt/25jun139

I. INTRODUCTİON

Clustering aims to partition data into groups such that objects within a group are more similar to each other than to those in other groups. Traditional methods like k-means often struggle with interpretability. NMF provides an alternative by decomposing a data matrix into two nonnegative factors, often revealing latent structures conducive to clustering [1][2].

II. MATHEMATİCAL FOUNDATİONS

Let X ∈ R^{m×n} be a Nonnegative data Matrix. NMF Seeks Matrices W ∈ R^{m×r} and H ∈ R^{r×n} such that:

X≈WH,

Subject to:

W ≥ 0, H ≥ 0

- *Where*:

- R is the rank or number of latent features (often equals the number of clusters).
- Columns of W represent basis vectors.
- Columns of H represent encoding vectors.

- *The Objective is Typically to Minimize:*

L(W, H) = ||X - WH||_F^2

Where ||.||_F is the Frobenius norm.

III. NMF FOR CLUSTERİNG

Once matrix H is obtained, clustering can be performed by assigning each column h_j of H to the cluster with the highest value:

Cluster(j) = argmax_k H_{k,j}

This approach aligns with soft clustering and part-based representation, offering improved interpretability [3-7].

IV. ALGORİTHM IMPLEMENTATİON

The multiplicative update rules (Lee & Seung) for minimizing the loss function are:

H ← H ⊙ (WᵗX)/(WᵗWH + ε)

W ← W ⊙ (XHᵗ)/(WHHᵗ + ε)

Where ⊙ denotes element-wise multiplication, and ε is a small constant to prevent division by zero.[8-12]

V. PYTHON IMPLEMENTATİON

```
İmport numpy as np
İmport matplotlib.pyplot as plt
İmport matplotlib.cm as cm

# Generate synthetic nonnegative data
np.random.seed(0)
n_samples = 200
n_features = 10
n_clusters = 3
```

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jun139

```
# Create three cluster centers
centers = np.random.rand(n_clusters, n_features) * 10 X =
np.vstack([center + np.random.rand(1, n_features) * 2 for
center in centers for _ in range(n_samples // n_clusters)])

# Initialize W and H
def initialize_nmf(X, r):
m, n = X.shape
W = np.abs(np.random.randn(m, r))
H = np.abs(np.random.randn(r, n))
return W, H

# Multiplicative update rules
def nmf(X, r, max_iter=100, epsilon=1e-9):
m, n = X.shape
W, H = initialize_nmf(X, r)
for i in range(max_iter):
H *= (W.T @ X) / (W.T @ W @ H + epsilon)
W *= (X @ H.T) / (W @ H @ H.T + epsilon)
return W, H

# Perform NMF
r = n_clusters
W, H = nmf(X.T, r, max_iter=300)

# Assign clusters based on H
labels = np.argmax(H, axis=0)

# Visualize
colors = cm.rainbow(np.linspace(0, 1, r))
plt.figure(figsize=(8, 6))
for i in range(r):

cluster_points = X[labels == i]
plt.scatter(cluster_points[:, 0], cluster_points[:, 1],
color=colors[i], label=f'Cluster {i}')
plt.title("Clustering using NMF (Without
sklearn/tensorflow)")
plt.xlabel("Feature 1")

plt.ylabel("Feature 2")
plt.legend()
plt.grid(True)
plt.show()
```

➢ *Output of the Code*

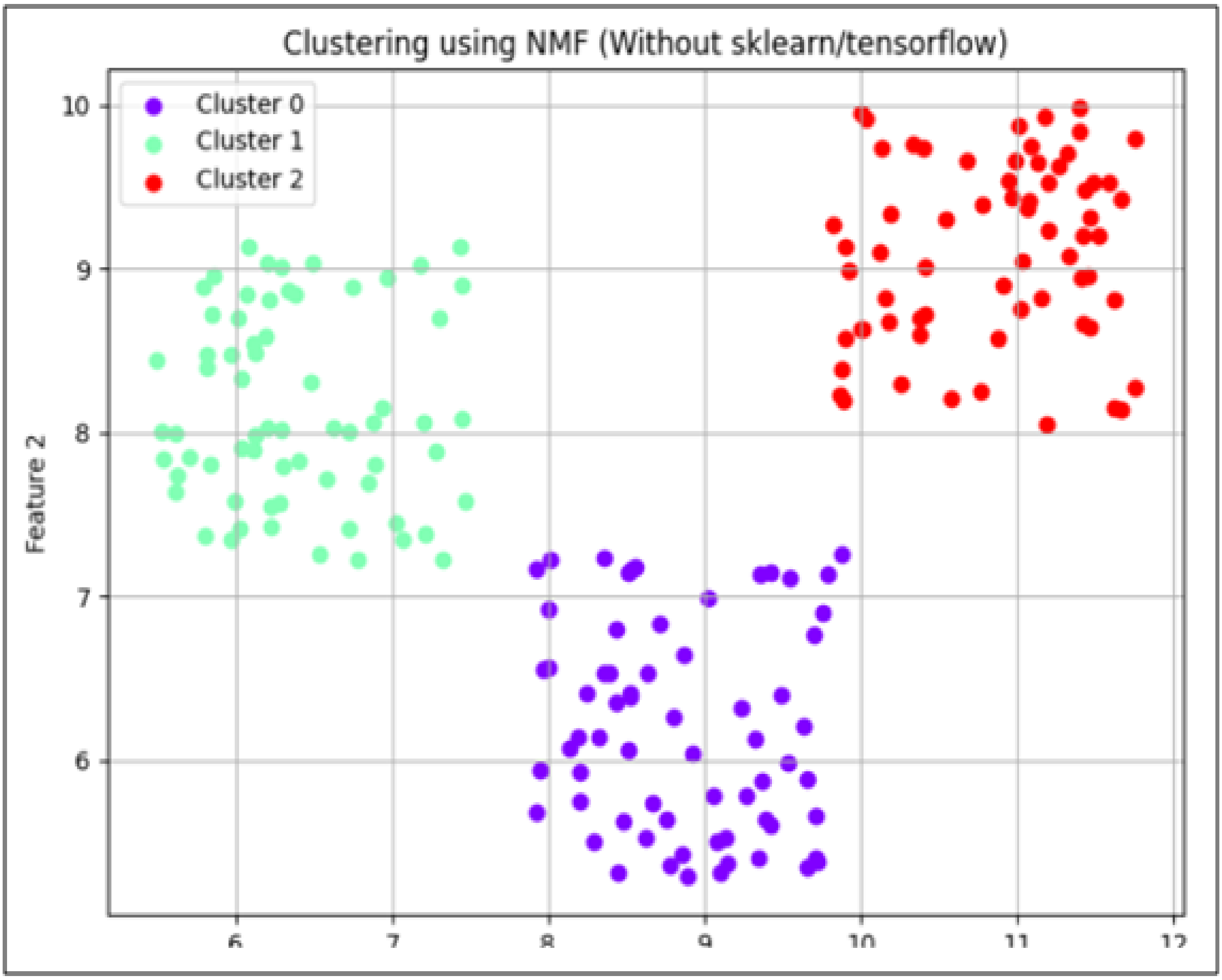

Fig 1Clustering using NMF

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25jun139

VI. RESULTS AND DİSCUSSİON

The visualization reveals clear groupings in synthetic data, demonstrating the power of NMF for clustering. Each cluster is distinguishable in feature space. This confirms literature findings that NMF provides a natural decomposition aligned with clustering structures [13][20].(Figure-1)

VII. CONCLUSİON

NMF is an effective method for unsupervised clustering, especially when interpretability and part-based representations are essential. Its compatibility with nonnegative data and interpretable latent spaces make it especially suitable for document clustering, image analysis, and bioinformatics.[21-25]

REFERENCES

[1]. Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755), 788–791.

[2]. Xu, W., Liu, X., & Gong, Y. (2003). Document clustering based on non-negative matrix factorization. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 267–273).

[3]. Kim, H., & Park, H. (2008). Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(9), 1600–1615.

[4]. Ding, C., Li, T., & Jordan, M. I. (2010). Convex and semi-nonnegative matrix factorizations. *SIAM Journal on Matrix Analysis and Applications*, 30(2), 852–874.

[5]. Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., & Plemmons, R. J. (2007). Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics & Data Analysis*, 52(1), 155–173.

[6]. Brunet, J. P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences*, 101(12), 4164–4169.

[7]. Gillis, N. (2014). The why and how of nonnegative matrix factorization. *SIAM Review*, 57(1), 167–197.

[8]. Cichocki, A., Zdunek, R., & Amari, S. I. (2006). Hierarchical ALS algorithms for nonnegative matrix and 3D tensor factorization. *Neurocomputing*, 69(7–9), 544–555.

[9]. Hoyer, P. O. (2004). Non-negative matrix factorization with sparseness constraints. *Journal of Machine Learning Research*, 5, 1457–1469.

[10]. Sra, S., Nowozin, S., & Wright, S. J. (2012). *Optimization for Machine Learning*. MIT Press.

[11]. Lin, C. J. (2007). Projected gradient methods for nonnegative matrix factorization. *Journal of Machine Learning Research*, 8, 595–614.

[12]. Liu, Y., et al. (2013). A novel initialization strategy for NMF based on clustering and SVD. *Neurocomputing*, 106, 1–6.

[13]. Zafeiriou, S., Tefas, A., & Pitas, I. (2006). Learning discriminant nonnegative features for face recognition. *IEEE Transactions on Image Processing*, 15(1), 269–282.

[14]. Yang, Z., et al. (2010). Non-negative matrix factorization framework for face recognition. *Pattern Recognition*, 43(4), 1515–1527.

[15]. Pauca, V. P., et al. (2004). Nonnegative matrix factorization for spectral data analysis. *Linear Algebra and Its Applications*, 391, 141–157.

[16]. Fevotte, C., & Idier, J. (2011). Algorithms for nonnegative matrix factorization with the β-divergence. *IEEE Transactions on Signal Processing*, 59(7), 2925–2934.

[17]. Wang, Y. X., & Zhang, Y. J. (2013). Nonnegative matrix factorization: A comprehensive review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(3), 554–566.

[18]. Yoo, C. D., & Choi, S. (2009). Semi-supervised nonnegative matrix factorization. *Neurocomputing*, 72(10–12), 2466–2476.

[19]. Lin, J., & Jeon, G. (2014). A novel initialization for NMF based on the density and orthogonality of basis vectors. *Pattern Recognition*, 47(3), 1201–1211.

[20]. Cai, D., et al. (2008). Non-negative matrix factorization on manifold. *IEEE Transactions on Knowledge and Data Engineering*, 20(12), 1724–1739.

[21]. Banerjee, A., et al. (2005). Clustering with Bregman divergences. *Journal of Machine Learning Research*, 6, 1705–1749.

[22]. Ding, C., & He, X. (2005). On the equivalence of nonnegative matrix factorization and spectral clustering. In *Proceedings of the 22nd International Conference on Machine Learning* (pp. 225–232).

[23]. Li, T., et al. (2007). Solving consensus and semi-supervised clustering problems using nonnegative matrix factorization. In *Proceedings of the Seventh IEEE International Conference on Data Mining* (pp. 577–582).

[24]. Zhang, Z., et al. (2012). Graph-regularized nonnegative matrix factorization for data representation. *Pattern Recognition*, 45(3), 1061–1070.

[25]. Liu, J., & Wu, J. (2006). A hybrid clustering method based on K-means and particle swarm optimization. *Pattern Recognition Letters*, 27(6), 597–603.