

1906003132015

Doğal Dil İşleme

BAİBÜ Bilgisayar Müh.

Dr. Öğr. Üyesi İsmail Hakkı Parlak

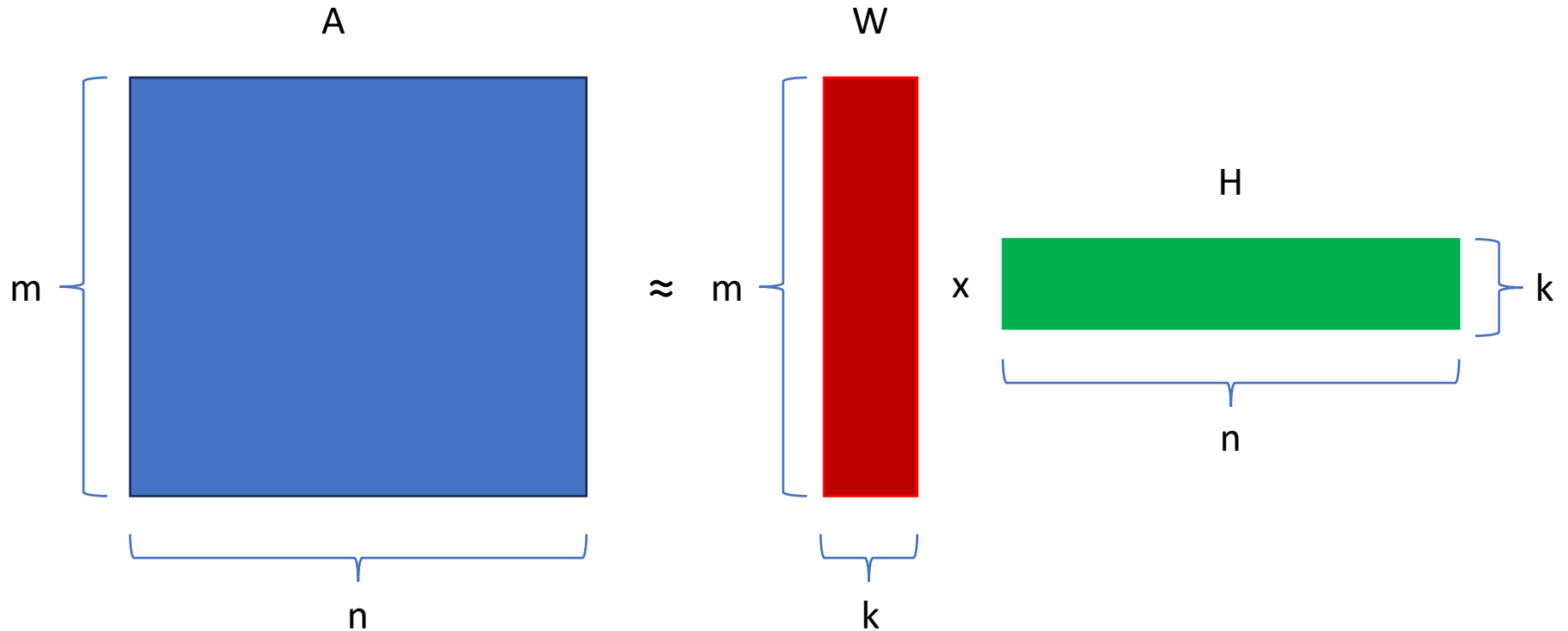
ismail.parlak@ibu.edu.tr

Oda: 335

Negatif Olmayan Matris Çarpanlarına Ayırma

- Negatif olmayan matris çarpanlarına ayırma (**N**on-**n**egative **m**atrix **f**actorization - NMF), $n \times m$ boyutlarındaki bir matrisi (A), $n \times k$ (W) ve $k \times m$ (H) boyutlarında 2 matrisin çarpımı haline getirmedir.
- $k \ll \min(m, n)$, $W \geq 0$, $H \geq 0$
- Elimizde bir A matrisi varken, $W \times H = A$ eşitliğini sağlayabilecek, elemanları negatif olmayan W ve H matrislerini direkt (analitik) hesaplamanın bir yolu yoktur. W ve H matrisleri nümerik yakınsama metotları ile yaklaşık olarak hesaplanabilir.

Negatif Olmayan Matris Çarpanlarına Ayırma



$m \times n$ kadar veri ile ifade edilen bir sistem, $(m \times k) + (k \times n) = k \times (m + n)$ kadar veri ile yaklaşık olarak ifade edilebilir hale gelir.

Negatif Olmayan Matris Çarpanlarına Ayırma

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 5 & 10 & 15 \\ \hline \end{array} = \begin{array}{|c|} \hline 1 \\ \hline 5 \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline \end{array} = \begin{array}{|c|} \hline 10 \\ \hline 50 \\ \hline \end{array} \times \begin{array}{|c|c|c|} \hline 0.1 & 0.2 & 0.3 \\ \hline \end{array}$$

$A \qquad W_1 \qquad H_1 \qquad W_2 \qquad H_2$

$\|A - W \times H\|^2$ diye gösterilen hatayı en aza indirgeyerek W ve H çarpanları elde edilir.

Ancak W ve H matrislerinin tüm elemanlarının ≥ 0 olması gerektiği unutulmamalıdır.

Tavsiye Sistemlerinde NMF

	F_1	F_2	F_3	F_4	F_5
K_1	?	2	?	2	0
K_2	3	2	1	?	?
K_3	?	4	2	2	?
K_4	?	2	?	2	?

K_i : Kullanıcılar

F_j : Filmler

A_{ij} : K_i kullanıcısının F_j filmine verdiği puan

Tavsiye Sistemlerinde NMF

	F ₁	F2	F3	F4	F5
K ₁	?	2	?	2	0
K ₂	3	2	1	?	?
K ₃	?	4	2	2	?
K ₄	?	2	?	2	?

 \approx

0	2
1	0
2	0
0	2

 \times

3	2	1	1	2
0	1	2	1	0

Tavsiye Sistemlerinde NMF

	F ₁	F2	F3	F4	F5
K ₁	0	2	4	2	0
K ₂	3	2	1	1	2
K ₃	6	4	2	2	4
K ₄	0	2	4	2	0

 \approx

0	2
1	0
2	0
0	2

 \times

3	2	1	1	2
0	1	2	1	0

H

W

Tavsiye Sistemlerinde NMF

	N_1	N_2
K_1	0	2
K_2	1	0
K_3	2	0
K_4	0	2

W

	F_1	F_2	F_3	F_4	F_5
N_1	3	2	1	1	2
N_2	0	1	2	1	0

H

W: Kullanıcıların film nitelikleri (N_1, N_2) ile ilişkileri

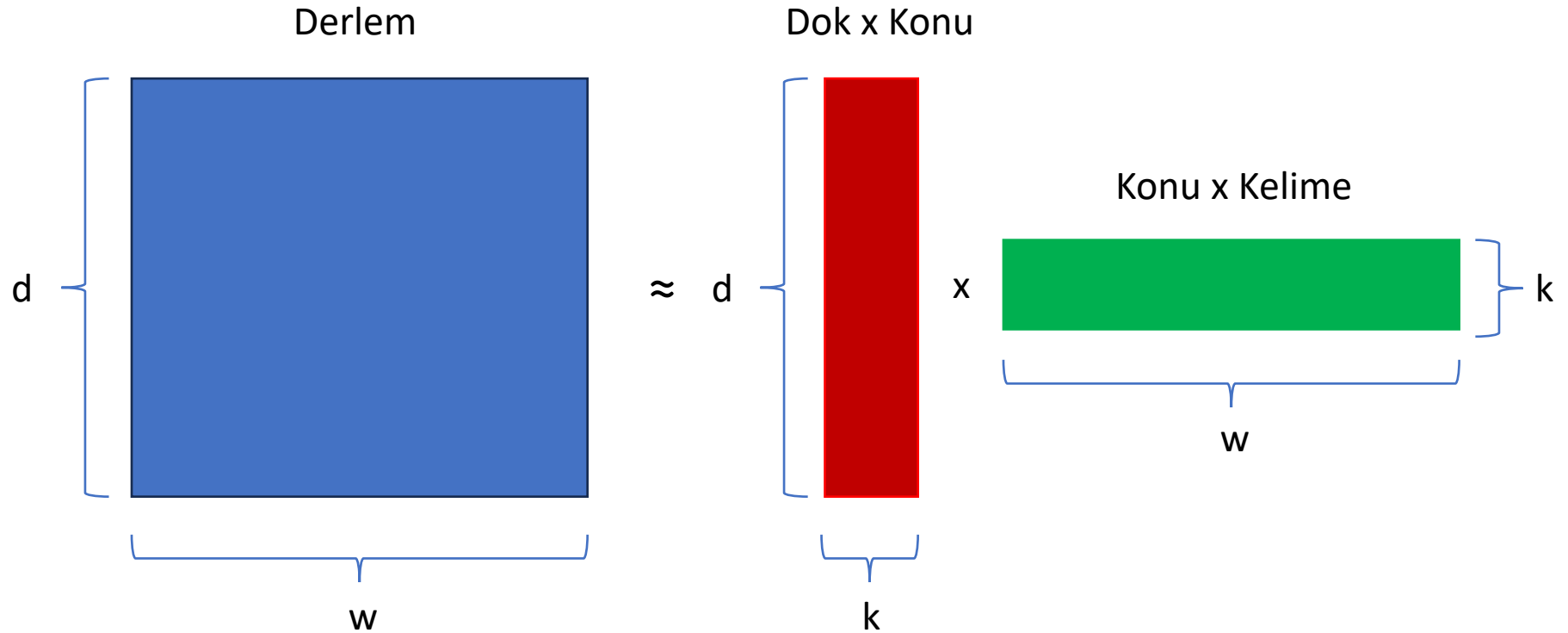
H: Filmlerin nitelik değerleri

N_i : Örtük nitelikler (Latent Features)

Konu Modelleme

- Konu modelleme (Topic modeling), derlem (corpus) içerisindeki metinlerde geçen konuları keşfedebilen bir kavramdır.
- Sınıflandırma algoritmalarının tersine, denetimsiz (unsupervised) olarak çalışır.
- Konu modelleme teknikleri ile bir dokümanlar havuzunda geçen konular keşfedilir ve havuzdaki dokümanlar bu otomatik keşfedilmiş konulara göre kümelendirilir (clustering).

Konu Modelleme



d : dokümanlar

w : kelime dağarcığı

k : konular