ISBN:1858-1382

J M S K

Jurnal Matematika, Statistika, & Komputasi
http://journal.unhas.ac.id/index.php/jmsk
Vol. 13, No. 2, 147-151, January 2017

Analisis Faktorisasi Matriks Tak Negatif

Andi Kresna Jaya†

Abstrak

Hasil perkalian matriks tak negatif W yang berukuran $n\times r$ dengan matriks tak negatif H yang berukuran $r\times m$ akan menghasilkan sebuah matriks tak negatif V berukuran $n\times m$, sedemikian sehingga dapat ditulis dalam bentuk persamaan matriks $W\times H = V$. Pada tulisan ini, kondisi seperti itu akan kita balik, yaitu dari sebuah matriks tak negatif V berukuran $n\times m$, akan dicari dua faktor matriks tak negatif yaitu matriks W yang berukuran $n\times r$ dengan matriks H yang berukuran $r\times m$, sedemikian sehingga berlaku $V=WH$. Pencarian tersebut mungkin tidak tunggal ataupun mungkin tidak ada, sehingga yang terjadi adalah hanya pendekatan terhadap matris V, yang didasarkan pada pertimbangan dimensi matriks-matriks faktornya dan norm matriks $V-WH$ adalah seminimum mungkin.

Kata kunci: *Faktorisasi*, *matriks tak negatif*, *dimensi*, *norm*.

1. Pendahuluan

Penelitian ini mengkaji aspek teoritis dari faktorisasi matriks tak negatif (fmtn) yang telah banyak digunakan oleh para peneliti di bidang terapan untuk menganalisis dan mereduksi dimensi gambar ataupun signal suara seperti yang telah dilakukan oleh Guillamet, Schiele dan Vitria, 2001 yang menggunakannya untuk klasifikasi gambar, Smaragdis dan Brown, 2003 [1] menggunakannya untuk transkripsi musik polyphonic, hal yang sama juga dikemukakan oleh Abdallah dan Plumbley pada tahun 2004 [2]. Hoyer di tahun 2002 [3] membahas pengkodean sparse tak negatif dan menghubungkannya dengan dekomposisi yang dilakukan pada sebuah matriks tak negatif.

Prinsip sederhana pada penggunaan fmtn setiap data asli suara atau gambar disusun dalam bentuk matriks V yang tak negatif berukuran $n\times m$. Kita akan mencari dua matriks W dan H yang ranknya adalah $r \le \min\{n,m\}$ dan hasil kali kedua matriks tersebut dapat mendekati keorisinalan matriks V. Setiap kolom dari matriks W merupakan vektor basis sedangkan matriks H menyatakan pembobotan yang diinginkan dapat mengaproksimasi kolom-kolom yang bersesuaian pada matriks V. Pencarian kedua matriks ini akan dilakukan berulang-ulang sehingga akan didapatkan hasil kali W dan H yang dapat mengaproksimasi V.

Sebuah Matriks berukuran $n\times m$ disebut tak negatif jika entri-entrinya adalah bilangan riil tak negatif dan disebut positif jika entri-entrinya adalah bilangan riil positif. Adapun karakteristik yang melekat dari bentuk matriks tak negatif ini jika ukurannya $n\times n$, adalah tracenya atau jumlah diagonal matriksnya tak negatif dan terdapat nilai eigen yang positif sedemikian hingga absolut nilai eigen yang lain tidak lebih besar dari nilai eigen tersebut, dan disebut nilai eigen maksimal [5]. Namun untuk matriks tak negatif $n\times m$ dapat dilihat pada nilai singular terbesarnya. Berikut ini untuk menentukan nilai skalar $V-WH$ akan digunakan bentuk nilai singularnya sebagai salah satu bentuk normnya. Norm ini berguna

† *Staf Pengajar pada Jurusan Matematika FMIPA Universitas Hasanuddin Makassar*

sebagai fungsi tujuan agar proses pencarian berulang-ulang atau iterasinya menjadi konvergen ke faktor-faktor yang dapat mendekati keorisinalitas matriks *V*.

2. Norm-norm untuk Matriks

Misalkan *A* adalah matriks berukuran $n\times m$ maka yang dimaksud dengan norm matriks adalah sebuah skalar yang menyatakan ukuran dari matriks yang bergantung pada entri-entrinya. Notasinya adalah $||A||$, sedemikian sehingga memenuhi sifat

- $||A|| \geq 0$ dan akan sama dengan nol jika $A = \mathbf{0}$.
- $||cA||=|c|\ ||A||$ untuk suatu c skalar
- $||A+B||\leq ||A|| + ||B||$ untuk matriks B berukuran $n\times m$

Bentuk norm matriks ini dapat diklasifikasikan atas

a. Jumlah nilai mutlak dari entri-entri pada baris yang terbesar, disebut norm-∞ dan dinotasikan oleh
$$\|A\|_\infty = \max_i \left\{ \sum_{j=1}^{m} |a_{ij}|, \quad i = 1,2,3,\cdots,n \right\}$$
b. Jumlah nilai mutlak entri-entri pada kolom yang terbesar, disebut norm-1 dan dinotasikan oleh
$$\|A\|_1 = \max_j \left\{ \sum_{i=1}^{n} |a_{ij}|, \quad j = 1,2,3,\cdots,m \right\}$$
c. Nilai singular terbesar dari A, dinotasikan oleh
$$\|A\|_{svd} = \sqrt{\lambda_{\max}\left(A^t A\right)} = \sqrt{\lambda_{\max}\left(AA^t\right)}$$
dengan $\lambda_{\max}(.)$ adalah nilai eigen terbesar.

d. Akar dari jumlah kuadrat tiap entri matriks disebut norm frobenius dan dinotasikan oleh bentuk
$$\|A\|_2 = \left(\sum_{ij} a_{ij}^{\ 2} \right)^{\frac{1}{2}}$$

Contoh 1:

Misalkan A matriks berukuran 2×3 berikut
$$A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$
Maka norm-normnya adalah
$$\|A\|_\infty = \max\{4,2\} = 4$$
$$\|A\|_1 = \max\{2,3,1\} = 3$$
$$\|A\|_{svd} = \sqrt{\lambda_{\max}\left(A^t A\right)} = \sqrt{\lambda_{\max}\left(AA^t\right)} = \sqrt{7.6056} = 2.7578$$
$$\|A\|_2 = \left(1^2 + 2^2 + 1^2 + 1^2 + 1^2 + 0^2\right)^{\frac{1}{2}} = 2\sqrt{2}$$

Masih banyak bentuk norm matriks yang lain, namun yang akan kita gunakan sebagai bentuk fungsi tujuan pada fmtn hanya keempat bentuk norm di atas.

3. Faktorisasi Matriks Tak Negatif

Untuk mendapatkan hampiran faktorisasi $V \approx WH$ adalah pemilihan dua matriks tak negatif sebagai faktor kemudian harus dilihat bentuk matriks $V-WH$ apakah normnya adalah yang paling kecil.

Contoh 2:
Misalkan $V=A$ pada contoh 1, kemudian kita pilih dua buah faktor matriks W dan H yang ranknya adalah $r=1 \leq \min\{2,3\}$ sebagai berikut

$$W = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ dan } H = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix},$$

maka bentuk norm matriks $V-WH$ adalah

$$\begin{aligned}
\|V - WH\|_{\infty} &= \max\{1,2\} = 2 \\
\|V - WH\|_{1} &= \max\{1,2,0\} = 2 \\
\|V - WH\|_{svd} &= \sqrt{\lambda_{\max}\left((V - WH)^t (V - WH)\right)} \\
&= \sqrt{\lambda_{\max}\left((V - WH)(V - WH)^t\right)} = \sqrt{2.6180} = 1.6180 \\
\|(V - WH)\|_{2} &= \left(0^2 + 1^2 + 0^2 + 1^2 + 1^2 + 0^2\right)^{\frac{1}{2}} = \sqrt{3}
\end{aligned}$$

Proses iterasi yang dilakukan untuk memperoleh norm minimum untuk matriks $V-WH$ dilakukan dengan cara menentukan terlebih dahulu dua buah faktor awal W_0 dan H_0 yang ranknya $r \leq \min\{n,m\}$, kemudian iterasi pertama akan menghasilkan dua buah faktor berikutnya dan seterusnya. Adapun iterasi yang dilakukan pada tiap entri-entri kedua faktor yang kita nyatakan dengan teorema berikut:

Teorema 1:
Norm matriks $V-WH$ merupakan barisan nilai-nilai yang monoton tak naik untuk bentuk iterasi entri-entri matriks faktor

$$w_{ik} \frac{\left(VH^T\right)_{ik}}{\left(WHH^T\right)_{ik}} \to w_{ik} \text{, untuk } i = 1,2,\cdots,n \text{ dan } k = 1,2,\cdots,r$$

dan

$$h_{kj} \frac{\left(W^T V\right)_{kj}}{\left(W^T WH\right)_{kj}} \to h_{kj} \text{, untuk } k = 1,2,\cdots,r \text{ dan } j = 1,2,\cdots,m,$$

asalkan $\left(WHH^T\right)_{ik} \neq 0$, untuk $i = 1,2,\cdots,n$ dan $k = 1,2,\cdots,r$ dan $\left(W^T WH\right)_{kj} \neq 0$, untuk $k = 1,2,\cdots,r$ dan $j = 1,2,\cdots,m$.

Contoh 3:
Untuk V pada contoh 2, misalkan matriks faktor awal diberikan oleh

$$W1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ dan } H1 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

Maka dengan menggunakan program dalam Matlab 7 diperoleh 3 iterasi pertama

$$W_2 = \begin{bmatrix} 1.3333 \\ 0.6667 \end{bmatrix} \text{ dan } H_2 = \begin{bmatrix} 1 & 1.5 & 0.5 \end{bmatrix};$$

$$W_3 = \begin{bmatrix} 1.2857 \\ 0.7143 \end{bmatrix} \text{ dan } H_3 = \begin{bmatrix} 0.9 & 1.5 & 0.6 \end{bmatrix};$$

$$W_4 = \begin{bmatrix} 1.3158 \\ 0.7018 \end{bmatrix} \text{ dan } H_4 = \begin{bmatrix} 0.9245 & 1.5189 & 0.5943 \end{bmatrix}$$

Nilai keempat norm untuk masing-masing faktor iterasi ditunjukkan pada tabel .

Tabel 1. Norm fmtn untuk 4 Iterasi Pertama

W dan H	$\lVert V-WH \rVert_1$	$\lVert V-WH \rVert_\infty$	$\lVert V-WH \rVert_{svd}$	$\lVert V-WH \rVert_2$
W_1,H_1	1	1	1	1.4142
W_2,H_2	0.6667	0.6667	0.6667	0.6667
W_3,H_3	0.6571	0.8571	0.6299	0.6312
W_4,H_4	0.6351	0.8342	0.6282	0.6293

Tabel di atas menunjukkan bahwa norm $\lVert V-WH \rVert_{svd}$ mempunyai pendekatan yang lebih baik, walaupun untuk kasus lain hal itu belum tentu benar. Kasus di atas juga memperlihatkan kepada kita bahwa iterasinya berjalan lambat untuk faktor awal yang diberikan. Namun dengan proses iterasi yang terus-menerus maka kita akan mendapatkan hampiran terbaik untuk fmtn.

4. Kesimpulan

Bentuk fmtn yang dapat dilakukan adalah meminimumkan norm $\lVert V-WH \rVert$ dan mengasumsikan norm sebagai bentuk fungsi tujuan, maka fungsi kendalanya diberikan adalah W dan H adalah positif.

Algoritma untuk faktorisasi matriks tak negatif ini adalah $h_{kj} \frac{\left(W^T V\right)_{kj}}{\left(W^T WH\right)_{kj}} \to h_{kj}$ dan $w_{ik} \frac{\left(VH^T\right)_{ik}}{\left(WHH^T\right)_{ik}} \to w_{ik}$ untuk meminimumkan norm.

Daftar Pustaka

[1] Judith C Brown. Paris Smaragdis. 2003. Non-Negative Factorization for Polyphonic Music Transcription, *IEEE Workshop on Application of Signal Proccessing to Audio and Acoustics*, pp 177 – 178.

[2] Mark D Plumbley. Samer A Abdallah, 2004. Polyphonic Music Transcription by Non-Negative Sparse Coding of Power Spectra, *Centre for Digital Music*, University of London.

[3] Patrik O Hoyer, 2002, Non-Negative Sparse Coding, *Proc. IEEE Workshop on Neural Networks for Signal Processing*.

[4] Hiroki Asari. 2004. Non-negative Matrix Factorization: A possible way to learn sound dictionaries. *Paper Tony Zador Lab*, Watson school of Biological Sciences cold Spring Harbor Laboratory.

[5] Andi Kresna Jaya. 2001. Masalah Nilai Karakteristik Invers Real Tak Negatif, *Tesis Magister*. PPs Magister Matematika, Institut Teknologi Bandung.