Using Nonnegative Matrix Factorization to Analyze a Set of Documents

Helena Šmigoc (University College Dublin)

Joint work with Barak Pearlmutter (Maynooth University)

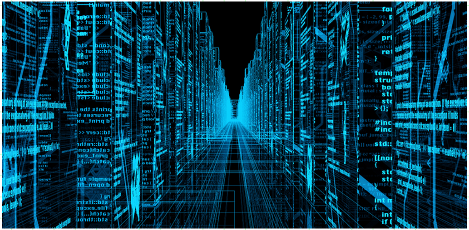
and Miao Wei (Maynooth University)

ILAS

Modules I teach

- Linear Algebra for Science, 1st module in linear algebra
- Mathematics of Google, 2nd module in linear algebra
- Applied Matrix Theory, MSc module

Why?



Factorization

$$6 = 3 \times 2$$

$$x^4 - 10x^3 + 35x^2 - 50x + 24 = (x-1)(x-2)(x-3)(x-4)$$

Matrix Factorization

$$A = \begin{pmatrix} 2 & -1 & 1 & 2 \\ -1 & 1 & -2 & -1 \\ 1 & -2 & 5 & 1 \\ 2 & -1 & 1 & 2 \end{pmatrix}$$

Matrix Factorization

$$A = \begin{pmatrix} 2 & -1 & 1 & 2 \\ -1 & 1 & -2 & -1 \\ 1 & -2 & 5 & 1 \\ 2 & -1 & 1 & 2 \end{pmatrix}$$

$$A = \frac{1}{21} \begin{pmatrix} 1 & 1 & -1 & 1 \\ -1 & 0 & 0 & 3 \\ 2 & -1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & -3 & 6 & 3 \\ 7 & 0 & -7 & 7 \\ -10 & 3 & 1 & 11 \\ 1 & 6 & 2 & 1 \end{pmatrix}$$

Matrix Factorization

$$A = \begin{pmatrix} 2 & -1 & 1 & 2 \\ -1 & 1 & -2 & -1 \\ 1 & -2 & 5 & 1 \\ 2 & -1 & 1 & 2 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 & 1 \\ 1 & 0 & -1 & 1 \end{pmatrix}$$

Can we find a factorization?

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 4 & 2 \\ 2 & 4 & 4 & 2 \\ 1 & 2 & 2 & 1 \end{pmatrix}.$$

- Factor this matrix as a product of a column and a row vector.

Approximate Factorization

$$A = \begin{pmatrix} 1 & 1 & 1 & 1.01 \\ 1 & 1 & 1.01 & 1 \\ 1 & 1.01 & 1 & 1 \\ 1.01 & 1 & 1 & 1 \end{pmatrix}.$$

- Can we find a matrix that is close to the matrix A that can be written as a product of a column by a row?

Approximate Factorization

$$A = \begin{pmatrix} 1 & 1 & 1 & 1.01 \\ 1 & 1 & 1.01 & 1 \\ 1 & 1.01 & 1 & 1 \\ 1.01 & 1 & 1 & 1 \end{pmatrix}.$$

$$A_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1.0025 & 1.0025 & 1.0025 & 1.0025 \end{pmatrix}$$

- Which solutions is better? What does it mean for a matrix B to be close to A?

Labeling Rows and Columns

	Alice	Becky	Cindy	Dora
Alien	4	1	4	5
Animal House	1	5	4	2
Beetlejuice	2	2	5	3
Jaws	5	1	5	5
Life of Brian	1	5	5	1

	Becky	Cindy	Dora	Alice
Animal House	5	4	2	1
Life of Brian	5	5	1	1
Beetlejuice	2	5	3	2
Jaws	1	5	5	5
Alien	1	4	5	4

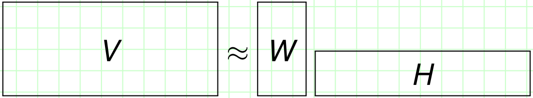
- Do the two tables contain the same information?
- What is the principle behind the orderings?

Nonnegative Matrix Factorization

Factors:

- non-negative
- of a given low rank r

$$\hat{V} = WH$$



r=2

	Alice	Becky	Cindy	Dora
Alien	4	1	4	5
Animal House	1	5	4	2
Beetlejuice	2	2	5	3
Jaws	5	1	5	5
Life of Brian	1	5	5	1

r=2

	Alice	Becky	Cindy	Dora
Alien	4	1	4	5
Animal House	1	5	4	2
Beetlejuice	2	2	5	3
Jaws	5	1	5	5
Life of Brian	1	5	5	1

$\rightarrow V$

$$V = \begin{pmatrix} 4 & 5 & 4 & 1 \\ 5 & 5 & 5 & 1 \\ 5 & 3 & 2 & 2 \\ 4 & 2 & 1 & 5 \\ 5 & 1 & 1 & 5 \end{pmatrix}$$

r=2

	Alice	Becky	Cindy	Dora
Alien	4	1	4	5
Animal House	1	5	4	2
Beetlejuice	2	2	5	3
Jaws	5	1	5	5
Life of Brian	1	5	5	1

$$W_2 = \begin{pmatrix} 7.6081 & 0 \\ 1.8627 & 5.6386 \\ 4.7191 & 2.6164 \\ 8.6624 & 0.090146 \\ 1.4702 & 6.3735 \end{pmatrix}$$

$$H_2 = \begin{pmatrix} 0.53813 & 0.10398 & 0.57089 & 0.61131 \\ 0 & 0.78183 & 0.61927 & 0.072416 \end{pmatrix}$$

$$\text{movies} \overbrace{\xrightarrow{W} \text{genres} \xrightarrow{H}}^{V} \text{people}$$

r=2

	Alice	Becky	Cindy	Dora
Alien	4	1	4	5
Animal House	1	5	4	2
Beetlejuice	2	2	5	3
Jaws	5	1	5	5
Life of Brian	1	5	5	1

$$V - W_2 H_2 = \begin{pmatrix} -0.094092 & 0.20891 & -0.34333 & 0.34911 \\ -0.0023903 & 0.39783 & -0.55524 & 0.45296 \\ -0.53945 & -0.53626 & 0.68572 & -0.074286 \\ 0.33855 & 0.028801 & -0.0010543 & -0.30193 \\ 0.20883 & -0.13592 & 0.21374 & -0.36031 \end{pmatrix}$$

Matrix Multiplication

$$\hat{v}_{\text{movie},\text{person}} = \sum_{g \in \text{genre}} w_{\text{movie},g}\, h_{g,\text{person}}$$

$$\begin{bmatrix} \hat{V} \end{bmatrix} = \begin{bmatrix} W \end{bmatrix} \begin{bmatrix} H \end{bmatrix}$$

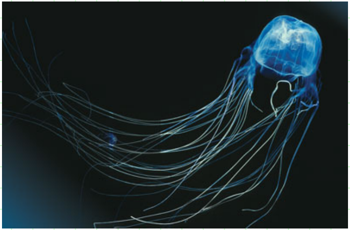
Matrix Multiplication

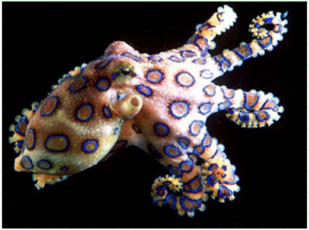
$$\begin{bmatrix} & & \\ & \hat{V} & \\ & & \end{bmatrix} = \sum_g \begin{bmatrix} \\ W_{\bullet,g} \\ \\ \end{bmatrix} \begin{bmatrix} & H_{g,\bullet} & \end{bmatrix}.$$

$$\begin{aligned} V &= \sum_{g=1}^{r} (g\text{-th column of } W) \cdot (g\text{-th row of } H) \\ &= \sum_{g \in \text{genres}} \text{Ratings Matrix for genre } g \\ &= \text{sum of rank-one per-genre matrices} \end{aligned}$$

Four Most Venomous Animals

Box Jellyfish, King Cobra, Marbled Cone Snail, Blue-Ringed Octopus



Term-by-Document Matrix of the Four Most Venomous Animals

		documents			
		jellyfish	cobra	snail	octopus
terms	venom	32	44	1	18
	death	9	3	0	2
	danger	6	4	0	4
	survive	2	0	0	1
	Madagascar	0	0	2	0

Matrix by a column vector

$$A = \begin{pmatrix} 32 & 44 & 1 & 18 \\ 9 & 3 & 0 & 2 \\ 6 & 4 & 0 & 4 \\ 2 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

- Determine the frequency of terms appearing in the first document, in the third document, in the first or third document.
- What is the frequency of terms in all the documents together?

Transpose

$$A^T = \begin{pmatrix} 32 & 9 & 6 & 2 & 0 \\ 44 & 3 & 4 & 0 & 0 \\ 1 & 0 & 0 & 0 & 2 \\ 18 & 2 & 4 & 1 & 0 \end{pmatrix}$$

- Which documents contain the third term, the fifth term, the third or the fifth term?

Bigger Example: Module descriptors

- Module descriptors are relatively short documents that give overviews of the courses.
- Corpus of documents: Module descriptors for 62 mathematics modules taught in the School of Mathematics and Statistics, University College Dublin
- All the words that appear in any of the 62 documents were collected. So-called stop words (common words like "and", "the", and the like), and all the words that appeared fewer than four times, were removed. Words were also down-cased and stemmed, so for example the terms *Eigenvalue* and *eigenvalues* are deemed equivalent.
- 290×62 matrix V of word counts.

Most Frequent Words	
word	**count**
function	117
theorem	75
linear	72
matrix	66
theory	58
equation	53
mathematical	52
mathematics	52
understand	49
science	48
problem	44

NMF

V is factored as

$$V \approx WH,$$

where W is an entry-wise non-negative $290 \times r$ matrix and H is an entry-wise nonnegative $r \times 62$ matrix.

W-matrix, $r = 2$

function	23.0
linear	11.1
matrix	11.0
equation	9.9
derivative	9.4
calculus	7.8
differential	6.5
solve	6.1
problem	5.9
mathematical	5.3
science	5.2
compute	5.1
variable	4.7
applications	4.7
integral	4.6

group	17.9
theorem	15.0
theory	6.9
ring	6.6
understand	4.0
structure	3.3
example	3.0
number	2.9
isomorphism	2.7
concepts	2.6
homomorphisms	2.5
sylow	2.3
subgroups	2.3
quotient	2.2
cauchy	2.2

W-matrix, $r = 3$

group	18.1
theorem	14.8
theory	6.8
ring	6.7
understand	3.9
structure	3.4
number	3.0
example	2.9
isomorphism	2.6
homomorphisms	2.6
concepts	2.5
sylow	2.3
subgroups	2.3
applications	2.2
quotient	2.2
cauchy	2.1
time	2.1
finite	2.1
algebraic	2.1
permitting	2.0

function	27.1
derivative	11.1
calculus	9.2
equation	6.3
differential	5.8
integral	5.7
problem	5.5
variable	5.3
graph	4.8
limit	4.6
solve	4.5
mathematics	4.3
calculate	3.9
applications	3.8
science	3.5
introduction	3.5
mathematical	3.5
method	3.4
polynomial	3.4
differentiation	3.3

matrix	19.1
linear	14.9
space	9.5
vector	8.7
algebra	7.1
basis	6.5
equation	6.4
compute	5.7
system	4.9
rank	3.5
complex	3.5
product	3.4
number	3.3
mathematical	3.3
science	3.3
solve	3.2
dimensional	3.0
eigenvalues	2.9
set	2.8
eigenvectors	2.7

H-matrix, $r = 3$

Module Names

Linear Algebra with Applications to Economics	Groups, Rings, and Fields	Differential Equations via Computer Algebra
0.0	0.3	0.0
0.0	0.0	0.1
0.2	0.0	0.0

NMF of ILAS 2017 Abstracts, $r = 2$, W matrix

Column 1		Column 2	
graph	**0.0324**	matrix	0.0579
eigenvalue	0.0156	matrice	0.0146
work	0.0117	**polynomial**	**0.0133**
operator	0.0116	problem	0.0117
set	0.0108	work	0.0093
result	0.0103	pattern	0.0088
bound	0.0100	algorithm	0.0078
positive	0.0097	complex	0.0076
number	0.0087	rank	0.0073
matrice	0.0081	arbitrary	0.0073
problem	0.0080	method	0.0072
spectral	0.0080	equation	0.0062
joint	0.0078	block	0.0060
quantum	0.0074	talk	0.0060
forcing	0.0072	positive	0.0059
space	0.0066	using	0.0059
linear	0.0063	joint	0.0059
zero	0.0053	spectrally	0.0058
tensor	0.0053	eigenvalue	0.0057
numerical	0.0053	numerical	0.0054
discus	0.0052	based	0.0054
state	0.0049	linear	0.0053
mean	0.0048	solution	0.0053

NMF of ILAS 2017 Abstracts, $r = 2$, H matrix

Row 1		Row 2	
suil_o	2.0341	peng-ruei_huang	2.0071
aida_abiad	1.9799	andrii_dmytryshyn	1.6782
siwaporn_mamart	1.9188	caiqin_song	1.5858
leonardo_de_lima	1.8832	little_hermie_monterde	1.5238
randy_davila	1.6149	jephian_c.-h._lin	1.4114
takashi_sano	1.4946	jillian_glassett	1.3730
chunli_deng	1.4837	robert_nowak	1.3507
willem_haemers	1.4830	raf_vandebril	1.2228
supalak_sumalroj	1.4707	hiroshi_nakazato	1.2224
michael_tait	1.4655	jongrak_lee	1.1917
sebastian_cioaba	1.3954	javier_perez_alvaro	1.1526
rachid_marsli	1.3578	yang_zhang	1.1496
carolyn_reinhart	1.3240	sasmita_barik	1.1209
veronika_furst	1.3011	richard_ferro	1.1052
gabriel_coutinho	1.2853	xingping_sheng	1.0834
changjiang_bu	1.2234	ravindra_bapat	1.0810
marko_orel	1.2162	hana_kim	1.0451
radu_balan	1.1748	ke_wei	1.0354
boris_brimkov	1.1201	marc_van_barel	1.0132
projesh_nath_choudhury	1.1042	kevin_vander_meulen	1.0096
masatoshi_ito	1.0501	judi_mcdonald	0.9713
domingos_cardoso	1.0329	joão_serra	0.9548
hugo_woerdeman	1.0254	ricardo_l._soto	0.9531
josh_carlson	1.0247	qing-wen_wang	0.9433
steve_butler	1.0176	shaun_fallat	0.9402
takeaki_yamazaki	1.0130	jim_cushing	0.9354

NMF of ILAS 2017 Abstracts, $r = 2$, H matrix

Column Judi		
judi_mcdonald	0.0012	0.9713

Judi McDonald, Spectrally Arbitrary Patterns over Different Fields

An $n \times n$ pattern is said to be spectrally arbitrary over a field F provided any n-th degree monic polynomial over F can be realized by a matrix with the given pattern and entries from F. In this talk we will look at algebraic and analytic properties of fields an illustrate how there properties can help us predict whether or not a pattern that is spectrally arbitrary over one field will be spectrally arbitrary over a subfield or field extension.

NMF of ILAS 2017 Abstracts, $r = 3$, W matrix

Column 1		Column 2		Column 3	
matrix	0.0701	graph	0.1185	positive	0.0174
eigenvalue	0.0203	eigenvalue	0.0284	matrice	0.0172
problem	0.0143	forcing	0.0177	work	0.0151
polynomial	0.0142	bound	0.0164	operator	0.0132
numerical	0.0115	number	0.0158	linear	0.0101
pattern	0.0114	spectral	0.0148	joint	0.0100
complex	0.0095	merging	0.0139	algebra	0.0097
range	0.0093	vertice	0.0125	problem	0.0094
arbitrary	0.0090	set	0.0124	result	0.0085
rank	0.0085	adjacency	0.0122	space	0.0079
method	0.0085	spectrum	0.0120	set	0.0072
algorithm	0.0081	bipartite	0.0105	talk	0.0068
block	0.0077	zero	0.0096	discus	0.0067
spectrally	0.0076	largest	0.0092	mean	0.0064
symmetric	0.0063	result	0.0092	quantum	0.0059
result	0.0060	connected	0.0086	matrices	0.0059
equation	0.0056	distance	0.0086	representation	0.0052
toeplitz	0.0056	classe	0.0086	definite	0.0050
using	0.0055	second	0.0078	map	0.0047
field	0.0055	quantum	0.0072	state	0.0045
entrie	0.0052	power	0.0072	using	0.0042
work	0.0051	determined	0.0068	bound	0.0042
new	0.0050	regular	0.0063	semidefinite	0.0041
function	0.0049	laplacian	0.0061	nonnegative	0.0040
				spectral	0.0039

NMF of ILAS 2017 Abstracts, $r = 3$, H matrix

Row 1		**Row 2**		**Row 3**	
peng-ruei_huang	1.8914	siwaporn_mamart	1.8966	projesh_nath_choudhury	2.0128
andrii_dmytryshyn	1.4312	leonardo_de_lima	1.3677	takashi_sano	1.9298
caiqin_song	1.3396	supalak_sumalroj	1.3073	harvey_blau	1.5343
jillian_glassett	1.2651	suil_o	1.0877	jin_liang	1.4936
jongrak_lee	1.2432	willem_haemers	1.0817	radu_balan	1.4724
jephian_c.-h._lin	1.1781	randy_davila	1.0085	minerva_catral	1.4042
robert_nowak	1.0246	aida_abiad	0.9561	mohammad_adm	1.4028
javier_perez_alvaro	1.0206	michael_tait	0.9449	rajendra_bhatia	1.3653
raf_vandebril	1.0176	marko_orel	0.8841	takeaki_yamazaki	1.3356
ravindra_bapat	0.9750	maría_robbiano	0.6490	masatoshi_ito	1.3136
sasmita_barik	0.9691	shahla_nasserasr	0.6278	alexander_belton	1.2738
marc_van_barel	0.9608	veronika_furst	0.6258	stephen_kirkland	1.1989
little_hermie_monterde	0.9445	sebastian_cioaba	0.5984	joshua_isralowitz	1.1739
hiroshi_nakazato	0.9345	carolyn_reinhart	0.5897	hugo_woerdeman	1.1628
ke_wei	0.9082	domingos_cardoso	0.5720	tanvi_jain	1.1103
xingping_sheng	0.9045	sarah_plosker	0.5559	miklós_pálfia	1.0887
anthony_austin	0.8959	chris_godsil	0.5240	lizhu_sun	1.0791
kevin_vander_meulen	0.8750	mahya_ghandehari	0.5140	man-duen_choi	1.0669
richard_ferro	0.8477	chassidy_bozeman	0.5072	dustin_mixon	1.0574
judi_mcdonald	0.8404	mary_flagg	0.4991	akaki_tikaradze	1.0360
patrick_rault	0.8383	enzo_wendler	0.4985	alicia_roca	1.0231
leonardo_de_lima	0.8254	gabriel_coutinho	0.4984	sivakumar_k.c.	1.0132
leonardo_robol	0.7920	xiaohong_zhang	0.4921	yongdo_lim	1.0099
ricardo_l._soto	0.7906	polona_oblak	0.4619	franklin_kenter	1.0095
brian_lins	0.7823	douglas_klein	0.4530	margarida_mitjana	1.0083
vasilije_perovic	0.7771	geir_dahl	0.4369	wai-shing_tang	0.9865
yang_zhang	0.7351	boris_brimkov	0.4219	dániel_virosztek	0.9830
tin-yau_tam	0.7154	enide_andrade	0.4092	jimmie_lawson	0.9780
hana_kim	0.7045	jiang_zhou	0.4038	simon_foucart	0.9770
philip_saltenberger	0.7027	chunli_deng	0.3848	adam_rutkowski	0.9762
jim_cushing	0.7015	josh_carlson	0.3760	cynthia_vinzant	0.9652

NMF of ILAS 2017 Abstracts, $r = 3$, H matrix

Column Judi

judi_mcdonald	0.8404	0.0001	0.0538

NMF of ILAS 2017 Abstracts, $r = 4$, W matrix

Column 1		Column 2		Column 3		Column 4	
graph	0.1203	thi	0.0489	positive	0.0550	matrix	0.0776
eigenvalue	0.0295	work	0.0174	operator	0.0514	polynomial	0.0186
forcing	0.0188	problem	0.0171	thi	0.0268	numerical	0.0144
number	0.0165	matrice	0.0117	space	0.0221	pattern	0.0137
bound	0.0155	algebra	0.0113	mean	0.0208	thi	0.0131
merging	0.0140	joint	0.0106	definite	0.0164	complex	0.0114
spectral	0.0139	linear	0.0098	result	0.0157	arbitrary	0.0114
adjacency	0.0131	eigenvalue	0.0095	matrice	0.0146	range	0.0106
spectrum	0.0128	tensor	0.0080	map	0.0140	eigenvalue	0.0101
vertice	0.0127	discus	0.0072	metric	0.0133	matrice	0.0100
set	0.0121	talk	0.0071	semidefinite	0.0132	spectrally	0.0096
thi	0.0110	representation	0.0066	real	0.0112	problem	0.0089
result	0.0110	quantum	0.0062	geometric	0.0106	method	0.0079
bipartite	0.0107	bound	0.0059	range	0.0081	equation	0.0078
zero	0.0099	method	0.0057	toeplitz	0.0074	algorithm	0.0078
matrix	0.0092	nonnegative	0.0057	classe	0.0072	rank	0.0075
distance	0.0091	state	0.0057	function	0.0072	block	0.0073
classe	0.0089	given	0.0054	numerical	0.0071	field	0.0068
connected	0.0088	algorithm	0.0050	shall	0.0068	using	0.0061
largest	0.0088	spectral	0.0047	work	0.0067	result	0.0059
power	0.0077	matrices	0.0046	set	0.0065	entrie	0.0055
second	0.0074	set	0.0046	relative	0.0063	symmetric	0.0055
laplacian	0.0067	using	0.0045	joint	0.0063	based	0.0055
determined	0.0067	application	0.0044	acting	0.0063	set	0.0054
upper	0.0064	rank	0.0042	power	0.0063	degree	0.0054
intersection	0.0064	result	0.0041	number	0.0061	solution	0.0051

NMF of ILAS 2017 Abstracts, $r = 4$, H matrix

Row 1		Row 2		Row 3		Row 4	
siwaporn_mamart	1.8689	harvey_blau	1.6828	takashi_sano	1.9014	peng-ruei_huang	1.8
leonardo_de_lima	1.4040	changjiang_bu	1.5133	takeaki_yamazaki	1.5935	andrii_dmytryshyn	1.4
supalak_sumalroj	1.2822	minerva_catral	1.4049	projesh_nath_choudhury	1.4740	jillian_glassett	1.3
willem_haemers	1.0800	franklin_kenter	1.3263	tanvi_jain	1.1343	caiqin_song	1.3
suil_o	1.0448	mohammad_adm	1.2600	jimmie_lawson	1.1112	little_hermie_monterde	1.0
randy_davila	1.0063	dustin_mixon	1.1710	dániel_virosztek	1.0923	jephian_c.-h._lin	1.0
michael_tait	0.9332	chunli_deng	1.1557	masatoshi_ito	0.9837	ravindra_bapat	0.9
aida_abiad	0.8998	roozbeh_gharakhloo	1.1487	miklós_pálfia	0.9480	hiroshi_nakazato	0.9
marko_orel	0.8730	lizhu_sun	1.1439	rajendra_bhatia	0.8899	raf_vandebril	0.9
maría_robbiano	0.6592	akaki_tikaradze	1.1367	martin_argerami	0.8872	xingping_sheng	0.9
shahla_nasserasr	0.6571	louis_deaett	1.0828	adam_rutkowski	0.8768	javier_perez_alvaro	0.9
veronika_furst	0.6093	radu_balan	1.0757	lajos_molnar	0.8623	judi_mcdonald	0.8
carolyn_reinhart	0.5886	margarida_mitjana	1.0664	joshua_isralowitz	0.8329	sasmita_barik	0.8
sebastian_cioaba	0.5645	sivakumar_k.c.	1.0608	jin_liang	0.8313	robert_nowak	0.8
domingos_cardoso	0.5480	stephen_kirkland	1.0584	douglas_farenick	0.7757	ke_wei	0.8
sarah_plosker	0.5452	alicia_roca	1.0553	farshid_abdollahi	0.7596	richard_ferro	0.8
chris_godsil	0.5121	rachid_marsli	1.0355	jongrak_lee	0.7435	jongrak_lee	0.8
chassidy_bozeman	0.5085	raphael_loewy	1.0219	m._cristina_câmara	0.7305	kevin_vander_meulen	0.8
enzo_wendler	0.5065	sebastian_cioaba	1.0127	fuad_kittaneh	0.6619	patrick_rault	0.8
mahya_ghandehari	0.5024	elias_jarlebring	1.0073	yongdo_lim	0.6541	marc_van_barel	0.7
mary_flagg	0.4819	gennadij_heidel	0.9829	césar_palencia	0.6524	yang_zhang	0.7
xiaohong_zhang	0.4675	aida_abiad	0.9800	eric_weber	0.6509	tin-yau_tam	0.7
polona_oblak	0.4563	stefan_güttel	0.9668	dariusz_chruscinski	0.5719	hana_kim	0.6
gabriel_coutinho	0.4533	silvia_marcaida	0.9619	man-duen_choi	0.5589	jim_cushing	0.6
douglas_klein	0.4379	mitsugu_hirasaka	0.9402	tianpei_jiang	0.5552	brian_lins	0.6
geir_dahl	0.4329	hugo_woerdeman	0.9220	radu_balan	0.5515	manami_chatterjee	0.6
jiang_zhou	0.4237	seung-hyeok_kye	0.8894	shaun_fallat	0.5306	zhongshan_li	0.6
boris_brimkov	0.4092	simon_telen	0.8883	aamena_alqabani	0.4990	leonardo_robol	0.6

NMF of ILAS 2017 Abstracts, $r = 4$, H matrix

Column Judi				
judi_mcdonald	0.0003	0.0678	0.0002	0.8690