

# ELEN 4720: Machine Learning for Signals, Information and Data

## Homework #4

Rakshith Kamath (rk3165)

December 12, 2021

### Problem-1 (Markov Chains)-35 Points

- 1) Use  $w_t$  to rank the teams by sorting in decreasing value according to this vector. List the top 25 team names (see accompanying file) and their corresponding values in  $w_t$  for  $t = 10, 100, 1000, 10000$ .

The top 25 teams at  $t=10, 100, 1000$  and  $10000$  are as shown below:

	t=10		t=100		t=1000		t=10000	
Rank	Team Name at t=10	Team Score at t=10	Team Name at t=100	Team Score at t=100	Team Name at t=1000	Team Score at t=1000	Team Name at t=10000	Team Score at t=10000
1	NorthCentral	0.01281	LSU	0.04364	LSU	0.08442	LSU	0.09137
2	LSU	0.01258	Wheaton	0.02582	Clemson	0.04553	Clemson	0.04926
3	Morningside	0.01151	NorthCentral	0.02444	OhioState	0.04345	OhioState	0.04701
4	Wheaton	0.01109	Clemson	0.0237	Georgia	0.03272	Georgia	0.03538
5	MinnSt-Mankato	0.00963	OhioState	0.02269	Alabama	0.02468	Alabama	0.0267
6	Muhlenberg	0.00951	Morningside	0.02262	Florida	0.0237	Florida	0.02562
7	OhioState	0.00924	Georgia	0.01716	Oregon	0.01927	Oregon	0.02078
8	UW-Whitewater	0.00917	StJohnsMN	0.01499	Auburn	0.01904	Auburn	0.02058
9	NorthDakotaSt	0.00914	WestFlorida	0.01337	Oklahoma	0.0187	Oklahoma	0.0202
10	Clemson	0.00894	Alabama	0.01292	PennState	0.01797	PennState	0.01943
11	WestFlorida	0.00878	UW-Whitewater	0.01284	Wisconsin	0.01516	Wisconsin	0.01639
12	MarianIN	0.00852	Florida	0.01244	NotreDame	0.01465	NotreDame	0.01582
13	StJohnsMN	0.00849	Muhlenberg	0.01124	Minnesota	0.0142	Minnesota	0.01535
14	MountUnion	0.00765	Oregon	0.01053	Baylor	0.0127	Baylor	0.01372
15	FerrisSt	0.00712	Auburn	0.01002	Michigan	0.01205	Michigan	0.01302
16	GrandView	0.00711	Oklahoma	0.01001	Utah	0.0117	Utah	0.01262
17	Georgia	0.00701	MinnSt-Mankato	0.00975	Memphis	0.01166	Memphis	0.01259
18	MaryHardin-Baylor	0.00678	MarianIN	0.00958	Iowa	0.01101	Iowa	0.01189
19	Lenoir-Rhyme	0.00616	PennState	0.00953	AppalachianSt	0.01092	AppalachianSt	0.01179
20	SlipperyRock	0.00609	ValdostaSt	0.00889	Texas	0.01019	Texas	0.011
21	Oregon	0.0059	Lenoir-Rhyme	0.00807	NorthDakotaSt	0.00922	Navy	0.00979
22	DelawareValley	0.00584	Wisconsin	0.00802	Navy	0.00907	NorthDakotaSt	0.00972
23	JamesMadison	0.00567	NotreDame	0.00787	TexasA&M	0.00893	TexasA&M	0.00966
24	ValdostaSt	0.00554	Minnesota	0.00753	BoiseSt	0.00831	BoiseSt	0.00894
25	Florida	0.00544	GrandView	0.00752	KansasSt	0.00733	KansasSt	0.00792

Table 1- Top 25 teams for different iterations

- 2) Plot the  $\|w_t - w_{inf}\|_1$  as a function of  $t$  for  $t = 1, 2, \dots, 10000$ .

The plot of first order norm between  $w_t$  and  $w_{inf}$  is shown below-

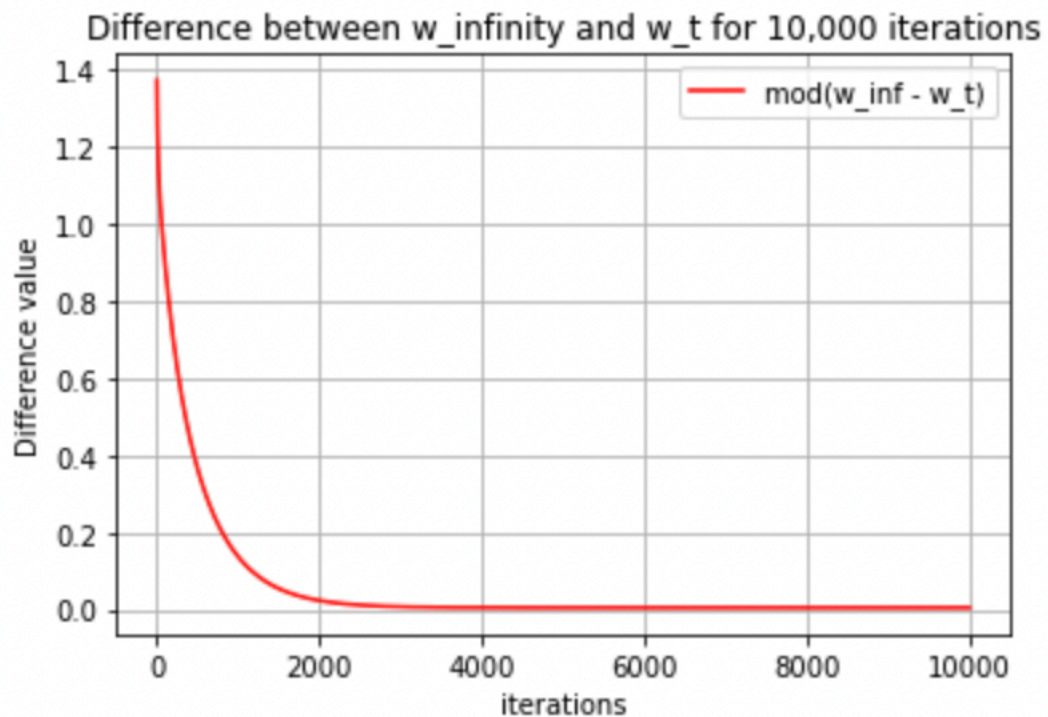


Figure 1 – Plot of difference between  $w_t$  and  $w_{\text{inf}}$  for various iterations

### **Problem-2 (Non-Negative Matrix Factorization)-40 Points**

- 1) Implement and run the NMF algorithm on this data using the divergence penalty. Set the rank to 25 and run for 100 iterations. This corresponds to learning 25 topics. Plot the objectives as a function of iteration.

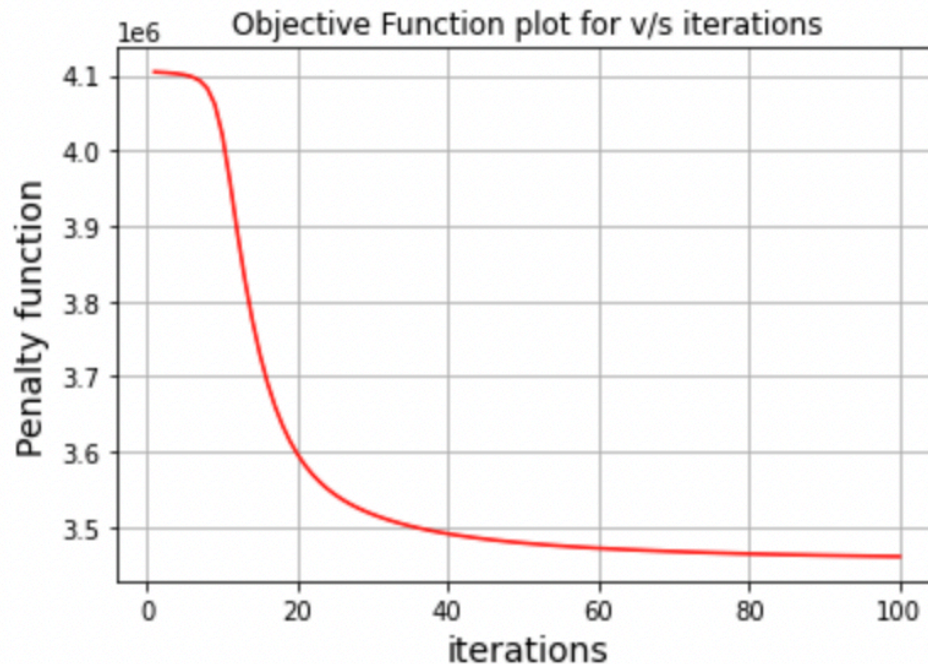


Figure 2- Plotting the objective functions for various iterations

- 2) After running the algorithm, normalize the columns of  $W$  so they sum to one. For each column of  $W$ , list the 10 words having the largest weight and show the weight. The  $i$ th row of  $W$  corresponds to the  $i$ th word in the dictionary provided in the data. Organize these lists in a 5x5 table. Each iteration and run gives different lists and one of them is shown below-

<b>List 1</b>	<b>List 2</b>	<b>List 3</b>	<b>List 4</b>	<b>List 5</b>
cause-0.009412740391593852	different-0.006196660903741392	space-0.005334491710036124	question-0.009596713419124252	star-0.007039061455934258
hospital-0.009778078175034228	possible-0.006291345876981601	area-0.005405629516275452	public-0.0096602098131097	sing-0.007169454797420469
care-0.010761380350949562	difficult-0.006364243516919029	large-0.006101703532858423	speak-0.010350769837364176	production-0.007690250398512398
treatment-0.011516611778384837	put-0.006693388989188376	small-0.00650268052862846	tell-0.010883436940062156	song-0.0080729158492686
patient-0.01237348639621335	change-0.006872275944537057	foot-0.0069563849357714745	reporter-0.012881048032032435	stage-0.008335033171264154
study-0.012547614842444908	mean-0.006913810395057294	build-0.007134186061199497	newspaper-0.015016209713450466	sound-0.009489362505741351
medical-0.013559135936461863	important-0.007857028614112227	design-0.007194143221718362	television-0.0154072983916745	performance-0.010460113256627024
doctor-0.015275769586859375	change-0.009290217439817516	mile-0.007951329443188087	article-0.01646490919918141	audience-0.012077444691404302
health-0.0161452433178594	question-0.00985191026791075	plant-0.012120205567094923	editor-0.021393772926729112	music-0.018490937449271397
drug-0.018426803947187573	problem-0.013141495253014729	water-0.013416951202534463	news-0.02319691718053052	play-0.01908561997256016
<b>List 6</b>	<b>List 7</b>	<b>List 8</b>	<b>List 9</b>	<b>List 10</b>
attack-0.00728473117359186	black-0.008502309940336343	pass-0.00759903255479587	democratic-0.010753525248647566	issue-0.01005662148530568
force-0.0095425225509898	teacher-0.009259641115159787	football-0.00788611178554505	leader-0.01162557097033255	agree-0.01049747072704876
leader-0.009993869015479204	member-0.0093863930456238	baseball-0.008112614246753526	candidate-0.012037960096685875	law-0.010685366137743655
country-0.010687954979035663	group-0.009628720279403874	league-0.009606599790440193	party-0.012263653572215328	approve-0.010782283351659087
states-0.011782538800905566	education-0.009656070411643054	coach-0.01498592482549283	state-0.012338749549111027	allow-0.011707351516440763
official-0.012194006729205321	college-0.01119787078188588	season-0.022431779577174706	election-0.013118107842647704	propose-0.01182585776627228
american-0.012575702696036938	class-0.0113388297523357	play-0.023625351150275163	republican-0.013551552124700843	require-0.01210136728190148
war-0.012963825237851749	child-0.011445509326881452	player-0.02523306902968124	campaign-0.01662178557171747	proposal-0.013331213050145152
government-0.013228458258446625	student-0.026606830916548425	game-0.0307484089608854	vote-0.0171704679252834	agreement-0.013426096097648499
military-0.013776405063089983	school-0.02969212360546027	team-0.03598436803385405	political-0.017483896558622338	plan-0.016038949560712903
<b>List 11</b>	<b>List 12</b>	<b>List 13</b>	<b>List 14</b>	<b>List 15</b>
arrest-0.008287276051245615	number-0.010433035640825246	neighborhood-0.0110037788840814	cook-0.006500045103730018	file-0.008252537758621673
night-0.008845491917820697	reduce-0.010517329666583851	town-0.011595203825830814	small-0.006586608288659227	trial-0.009134080402745775
shoot-0.009364499279690486	system-0.010682523310133916	live-0.012035481251161247	fresh-0.0069906432184353344	legal-0.009580707194817111
street-0.009687016980974497	pay-0.01086181770064104	build-0.012529609162937546	add-0.007151425304081844	investigation-0.009859409549010087
fire-0.010975304559946742	money-0.011146479553322505	area-0.01414037060336219	white-0.007526362790884055	judge-0.013012467985415908
officer-0.013202848224507244	cost-0.01439065104484476	resident-0.015525783307140556	red-0.007928506904443891	law-0.014344598342478778
man-0.015560730346009253	increase-0.01536916198404507	building-0.015693386233901405	taste-0.00830875495226607	charge-0.015831587337500824
kill-0.015590802352675228	state-0.015413304301848268	community-0.01614155441215556	serve-0.009256963045747477	court-0.017452983421392503
car-0.015877358014822263	program-0.01923216231000436	local-0.016404626674257364	restaurant-0.00970076593568041	lawyer-0.017970583609028798
police-0.02787703221323473	percent-0.019272595075074412	city-0.03387030546093585	food-0.01388256622142715	case-0.020757239362975157
<b>List 16</b>	<b>List 17</b>	<b>List 18</b>	<b>List 19</b>	<b>List 20</b>
friend-0.006700846507191648	trade-0.010512729760013328	present-0.006643708999456104	firm-0.01261670729095129	receive-0.019339104460246143
wear-0.006984198125080184	rise-0.010695678461406436	photograph-0.006852111959798892	bank-0.013159337522211295	wife-0.01955683645003314
life-0.007292753457351808	government-0.01102531043826805	image-0.006909845757406222	chairman-0.014501869495864555	marry-0.02439267687491522
man-0.007495550481608128	rate-0.01253734844528647	exhibition-0.007230479059792089	pay-0.01559745024799095	family-0.026707287963848428
little-0.007649594144209969	percent-0.012811766579299472	collection-0.007753092932790086	chief-0.015948033374572637	graduate-0.034991439292747395
lot-0.0076764542498847383	american-0.013747686480516228	movie-0.007772418231552851	financial-0.016747989041867705	daughter-0.0358252977239537
ask-0.007936738509429478	country-0.014880668717941437	museum-0.007778594437982067	president-0.01932954957833099	mrs-0.043323615856516236
feel-0.008280690017363665	states-0.01706808698030274	artist-0.012096858515905405	business-0.020654732791094768	son-0.04600819889090318
tell-0.009012599930387083	economy-0.017318332842696033	film-0.012448107821820402	executive-0.029252204696305833	mother-0.046193320845499465
thing-0.011012508852528146	economic-0.018221665186383434	art-0.018992605401853033	company-0.03658656748332372	father-0.05517280993384511
<b>List 21</b>	<b>List 22</b>	<b>List 23</b>	<b>List 24</b>	<b>List 25</b>
word-0.005684406740174328	game-0.010013266036434113	site-0.007646939819645493	investor-0.012647385968210413	sign-0.00924756164450044
page-0.005927700742897245	lose-0.010376882131888667	product-0.007692728238184526	analyst-0.013967384650570363	employee-0.011121360579501484
author-0.006434000922546304	final-0.010396701713512556	information-0.00840600704404530	company-0.017411994438471586	manager-0.011409635230115027
world-0.006596280097988507	hit-0.011277215682222879	technology-0.008461700794365899	sale-0.01745152281582083	official-0.012212711081465655
story-0.006755581406700477	play-0.01155546255095985	service-0.009386549621974057	sell-0.01839763623051942	worker-0.013127740875418414
history-0.007475601719546718	score-0.013362898116819136	sell-0.009545429687156774	price-0.019386174454492453	union-0.014622776885159327
man-0.009190887873700837	victory-0.014641790043768577	customer-0.009922618203670448	stock-0.020475466103393538	announce-0.016894514911662602
life-0.01155849101507897	third-0.015067008994144909	store-0.010289079882486327	share-0.020515749321143285	job-0.02050250800664884
book-0.014683825455297492	second-0.02085368805537199	computer-0.016059295642096877	market-0.022873225506400918	contract-0.02368874743575929
write-0.014830764819576266	win-0.02647614516975707	company-0.018596390710011866	percent-0.02746687511321618	yesterday-0.026321310149430215