

# HW5-Report

Wenbo Song  
ws2505

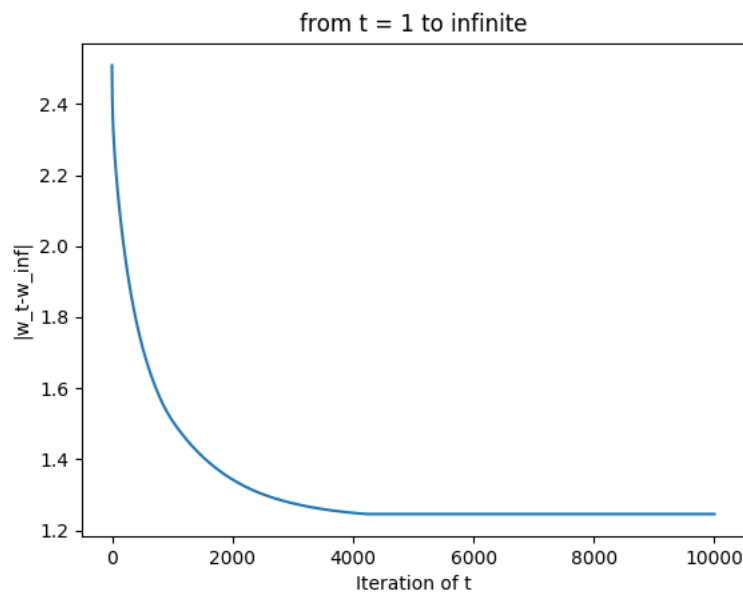
## Problem 1

a ) The top 25 teams and their corresponding values in wt for  $t = 10, 100, 1000, 10000$  are

1	MountUnion	0.018356136	MountUnion	0.065704577	Alabama	0.042282267	Alabama	0.0511138
2	MaryHardin-Baylor	0.012997452	Alabama	0.022962693	Georgia	0.034923104	Georgia	0.042176138
3	StFrancisIN	0.011045059	Georgia	0.019236704	MountUnion	0.034023506	OhioState	0.033563483
4	UW-Oshkosh	0.009959058	UW-Oshkosh	0.017905183	OhioState	0.027730637	Clemson	0.032494066
5	TAMU-Commerce	0.009294898	OhioState	0.014858506	Clemson	0.026893888	Oklahoma	0.031654142
6	BrockportSt	0.009168531	Clemson	0.014718529	Oklahoma	0.026177457	Wisconsin	0.031089781
7	NorthDakotaSt	0.009007857	Oklahoma	0.014204647	Wisconsin	0.025686303	CentralFlorida	0.028809804
8	Alabama	0.008993574	Wisconsin	0.013756421	CentralFlorida	0.02384565	Auburn	0.026505833
9	Morningside	0.008794384	CentralFlorida	0.013051149	Auburn	0.021932937	PennState	0.024029384
10	Georgia	0.008379974	MaryHardin-Baylor	0.012608741	PennState	0.019862347	NotreDame	0.018049514
11	JamesMadison	0.008075573	TAMU-Commerce	0.012499088	NotreDame	0.014938245	TCU	0.017211902
12	DelawareValley	0.00793908	Auburn	0.011958752	TCU	0.014246525	MiamiFL	0.014376984
13	Wartburg	0.007578921	StFrancisIN	0.011093594	MiamiFL	0.011902924	MichiganSt	0.01423978
14	IndianaPA	0.007222175	PennState	0.010702206	MichiganSt	0.011772289	OklahomaSt	0.013635968
15	Clemson	0.007081149	BrockportSt	0.009478007	OklahomaSt	0.011288873	Iowa	0.012912683
16	MinnSt-Mankato	0.006954954	NotreDame	0.008178957	Iowa	0.010673526	Northwestern	0.012504788
17	CentralFlorida	0.00675317	DelawareValley	0.007976053	Northwestern	0.010334466	IowaSt	0.012370087
18	OhioState	0.006671281	NorthDakotaSt	0.007953492	IowaSt	0.010248423	LSU	0.01156465
19	Wisconsin	0.006636422	FrostburgSt	0.007860816	LSU	0.00957317	SouthernCal	0.011438819
20	Reinhardt	0.006579947	TCU	0.007818115	NorthDakotaSt	0.009492203	Washington	0.011191002
21	StThomas	0.006016489	MinnSt-Mankato	0.007645491	SouthernCal	0.009478463	NorthDakotaSt	0.010939289
22	Oklahoma	0.005948038	CaseWesternReserve	0.007156142	Washington	0.009284822	Stanford	0.010548805
23	FrostburgSt	0.005506445	JohnCarroll	0.006855438	UW-Oshkosh	0.009237102	MississippiSt	0.010317239
24	Ashland	0.005367864	MiamiFL	0.00654646	Stanford	0.008742289	VirginiaTech	0.010060003
25	Assumption	0.005324477	Harding	0.006451953	MississippiSt	0.008541334	NorthCarolinaSt	0.009595348

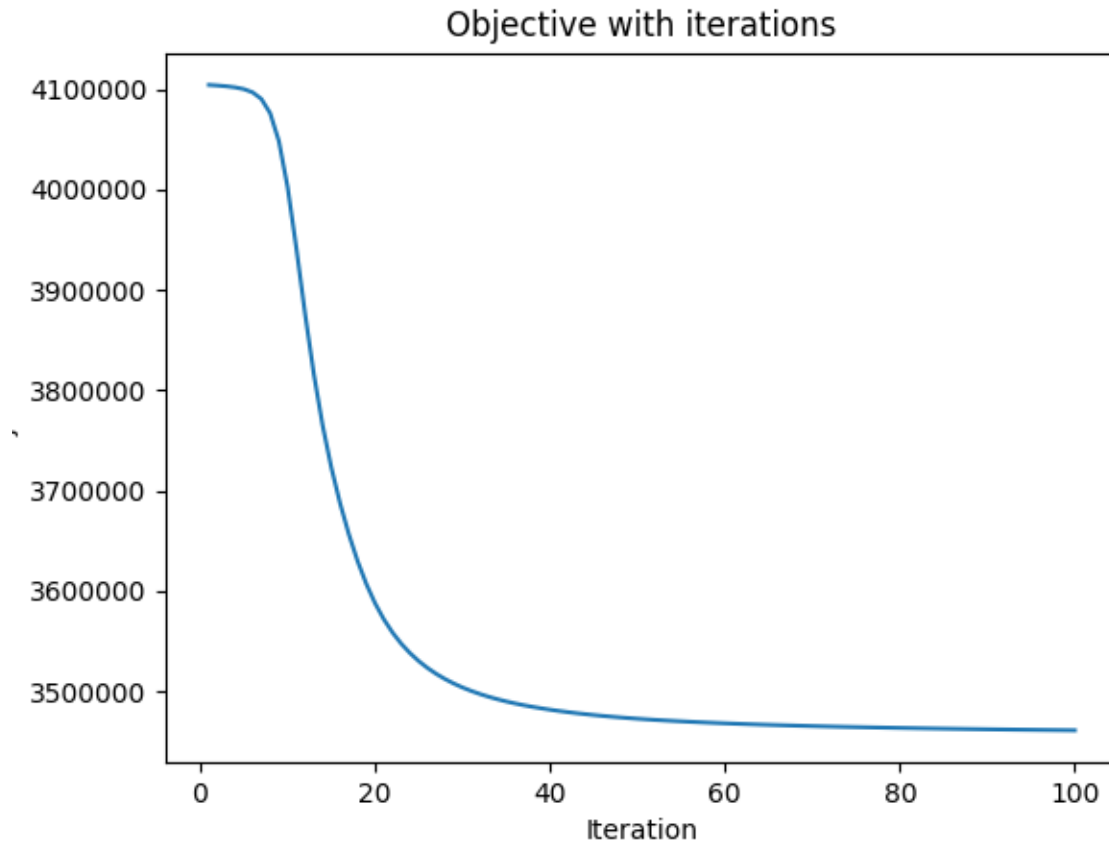
b )

Plot  $\|w_t - w_{\infty}\|$  as a function of  $t$  for  $t = 1, \dots, 10000$ .



## Problem 2

a ) Implement and run the NMF algorithm on this data using the divergence penalty. Plot the objective as a function of iteration.



b ) 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 with the data. Organize these lists in a  $5 \times 5$  table.

Topic_1	Topic_2	Topic_3	Topic_4	Topic_5
('0.034', 'company')	('0.021', 'political')	('0.022', 'food')	('0.018', 'night')	('0.007', 'view')
('0.026', 'executive')	('0.014', 'leader')	('0.011', 'serve')	('0.016', 'visit')	('0.007', 'question')
('0.019', 'president')	('0.013', 'campaign')	('0.010', 'restaurant')	('0.015', 'hour')	('0.006', 'idea')
('0.019', 'business')	('0.012', 'vote')	('0.009', 'taste')	('0.012', 'morning')	('0.006', 'great')
('0.013', 'industry')	('0.011', 'election')	('0.009', 'fresh')	('0.011', 'weekend')	('0.006', 'world')
('0.013', 'chief')	('0.011', 'party')	('0.008', 'cook')	('0.011', 'trip')	('0.006', 'sense')
('0.009', 'chairman')	('0.010', 'candidate')	('0.008', 'eat')	('0.009', 'open')	('0.006', 'mean')
('0.009', 'vice')	('0.009', 'republican')	('0.008', 'add')	('0.009', 'walk')	('0.005', 'point')
('0.009', 'director')	('0.009', 'support')	('0.007', 'pound')	('0.009', 'start')	('0.005', 'fact')
('0.009', 'network')	('0.009', 'democratic')	('0.007', 'chicken')	('0.009', 'home')	('0.005', 'different')
Topic_6	Topic_7	Topic_8	Topic_9	Topic_10
('0.016', 'car')	('0.014', 'wear')	('0.044', 'school')	('0.014', 'military')	('0.025', 'percent')
('0.013', 'train')	('0.011', 'white')	('0.029', 'student')	('0.013', 'war')	('0.016', 'price')
('0.012', 'driver')	('0.009', 'man')	('0.015', 'class')	('0.013', 'government')	('0.015', 'market')
('0.011', 'official')	('0.009', 'black')	('0.014', 'college')	('0.013', 'states')	('0.012', 'stock')
('0.011', 'plane')	('0.008', 'eye')	('0.013', 'education')	('0.013', 'american')	('0.011', 'rate')
('0.010', 'flight')	('0.008', 'dress')	('0.013', 'child')	('0.012', 'force')	('0.010', 'rise')
('0.010', 'vehicle')	('0.008', 'stand')	('0.012', 'program')	('0.011', 'official')	('0.010', 'share')
('0.010', 'traffic')	('0.007', 'little')	('0.011', 'teacher')	('0.010', 'attack')	('0.009', 'company')
('0.009', 'fly')	('0.007', 'blue')	('0.010', 'parent')	('0.010', 'country')	('0.009', 'sell')
('0.009', 'truck')	('0.007', 'woman')	('0.009', 'community')	('0.009', 'kill')	('0.008', 'fall')
Topic_11	Topic_12	Topic_13	Topic_14	Topic_15
('0.039', 'book')	('0.019', 'life')	('0.018', 'law')	('0.018', 'computer')	('0.021', 'music')
('0.026', 'write')	('0.013', 'child')	('0.016', 'rule')	('0.013', 'technology')	('0.018', 'play')
('0.014', 'writer')	('0.013', 'man')	('0.014', 'court')	('0.013', 'system')	('0.013', 'performance')
('0.013', 'publish')	('0.013', 'woman')	('0.012', 'case')	('0.010', 'company')	('0.011', 'audience')
('0.013', 'editor')	('0.013', 'family')	('0.011', 'legal')	('0.010', 'information')	('0.009', 'stage')
('0.012', 'author')	('0.012', 'young')	('0.011', 'lawyer')	('0.009', 'available')	('0.009', 'theater')
('0.011', 'page')	('0.012', 'tell')	('0.010', 'decision')	('0.008', 'cost')	('0.009', 'film')
('0.011', 'novel')	('0.011', 'friend')	('0.010', 'file')	('0.008', 'internet')	('0.009', 'production')
('0.010', 'read')	('0.011', 'thing')	('0.009', 'state')	('0.007', 'machine')	('0.009', 'dance')
('0.010', 'story')	('0.009', 'love')	('0.009', 'issue')	('0.007', 'number')	('0.009', 'perform')
Topic_16	Topic_17	Topic_18	Topic_19	Topic_20
('0.026', 'building')	('0.018', 'game')	('0.013', 'health')	('0.015', 'news')	('0.022', 'charge')
('0.024', 'city')	('0.017', 'season')	('0.013', 'study')	('0.011', 'question')	('0.018', 'police')
('0.016', 'resident')	('0.014', 'team')	('0.013', 'drug')	('0.011', 'editor')	('0.016', 'case')
('0.016', 'house')	('0.013', 'baseball')	('0.011', 'report')	('0.011', 'ask')	('0.013', 'investigation')
('0.014', 'build')	('0.013', 'hit')	('0.011', 'percent')	('0.011', 'issue')	('0.012', 'lawyer')
('0.012', 'area')	('0.013', 'player')	('0.010', 'doctor')	('0.009', 'tell')	('0.011', 'trial')
('0.011', 'neighborhood')	('0.011', 'play')	('0.010', 'medical')	('0.008', 'meeting')	('0.010', 'crime')
('0.011', 'live')	('0.010', 'start')	('0.010', 'patient')	('0.008', 'statement')	('0.010', 'criminal')
('0.010', 'community')	('0.010', 'contract')	('0.009', 'research')	('0.008', 'article')	('0.010', 'officer')
('0.010', 'street')	('0.009', 'pitch')	('0.008', 'cause')	('0.008', 'official')	('0.010', 'official')

Topic_21	Topic_22	Topic_23	Topic_24	Topic_25
('0.025', 'art')	('0.018', 'water')	('0.019', 'play')	('0.014', 'money')	('0.052', 'mrs')
('0.015', 'artist')	('0.014', 'plant')	('0.019', 'game')	('0.014', 'state')	('0.047', 'father')
('0.012', 'museum')	('0.008', 'mile')	('0.017', 'team')	('0.013', 'budget')	('0.041', 'son')
('0.012', 'photograph')	('0.007', 'area')	('0.017', 'win')	('0.013', 'bill')	('0.034', 'graduate')
('0.010', 'exhibition')	('0.007', 'place')	('0.013', 'second')	('0.012', 'pay')	('0.033', 'retire')
('0.009', 'design')	('0.006', 'tree')	('0.012', 'player')	('0.012', 'program')	('0.031', 'daughter')
('0.009', 'collection')	('0.006', 'clean')	('0.011', 'point')	('0.012', 'plan')	('0.025', 'receive')
('0.009', 'image')	('0.006', 'land')	('0.010', 'season')	('0.011', 'propose')	('0.024', 'president')
('0.009', 'picture')	('0.006', 'farm')	('0.010', 'victory')	('0.011', 'tax')	('0.024', 'mother')
('0.009', 'painting')	('0.006', 'grow')	('0.010', 'coach')	('0.011', 'proposal')	('0.023', 'marry')