CS6220 Data Mining Fall 2014 Homework 4, Wei Luo

1. Time Series

1.1 X=[71 73 80 80 80 78 76 75 73 71 71 71 73 75 76 76 68 76 76 75] Y=[69 69 73 79 80 79 78 76 73 72 71 70 70 69 69 69 71 73 75 76]

L_1 *norm* : $D(X,Y) = \sum_i |X_i - Y_i|$=51

1.2 Using the theory of dynamic programming, we can form the following table:

Y \ X	71	73	80	80	80	78	76	75	73	71	71	71	73	75	76	76	68	76	76	75
69	2	6	17	28	39	48	55	61	65	67	69	71	75	81	88	95	96	103	110	116
69	4	6	17	28	39	48	55	61	65	67	69	71	75	81	88	95	96	103	110	116
73	6	4	11	18	25	30	33	35	35	37	39	41	41	43	46	49	54	57	60	62
79	14	10	5	6	7	8	11	15	21	29	37	45	47	45	46	49	60	57	60	64
80	23	17	5	5	5	7	11	16	22	30	38	46	52	50	49	50	61	61	61	65
79	31	23	6	6	6	6	9	13	19	27	35	43	49	53	52	52	61	64	64	65
78	38	28	8	8	8	6	8	11	16	23	30	37	42	45	47	49	59	61	63	66
76	43	31	12	12	12	8	6	7	10	15	20	25	28	29	29	29	37	37	37	38
73	45	31	19	19	19	13	9	8	7	9	11	13	13	15	18	21	26	29	32	34
72	46	32	27	27	27	19	13	11	8	8	9	10	11	14	18	22	25	29	33	35
71	46	34	36	36	36	26	18	15	10	8	8	8	10	14	19	23	25	30	34	37
70	47	37	44	46	46	34	24	20	13	9	9	9	11	15	20	25	25	31	36	39
70	48	40	47	54	56	42	30	25	16	10	10	10	12	16	21	26	27	31	37	41
69	50	44	51	58	65	51	37	31	20	12	12	12	14	18	23	28	27	34	38	43
69	52	48	55	62	69	60	44	37	24	14	14	14	16	20	25	30	28	34	41	44
69	54	52	59	66	73	69	51	43	28	16	16	16	18	22	27	32	29	35	41	47
71	54	54	61	68	75	76	56	47	30	16	16	16	18	22	27	32	32	34	39	43
73	56	54	61	68	75	80	59	49	30	18	18	18	16	18	21	24	29	32	35	37
75	60	56	59	64	69	72	60	49	32	22	22	22	18	16	17	18	25	26	27	27
76	65	59	60	63	67	69	60	50	35	27	27	27	21	17	16	16	24	24	24	25

So $DTW(X,Y) = 25$, and the optimal warping path is the path shown in red in the table.

2. Graph

2.1 The adjacency matrix A for G is:

	1	2	3	4	5	6
1	0	0	1	1	0	0
2	0	0	1	0	0	0
3	0	1	0	1	1	1
4	1	0	0	1	0	0
5	0	0	0	0	0	1
6	0	1	0	0	0	0

2.2 The column stochastic matrix M for G is:

	1	2	3	4	5	6
1	0	0	0	1/2	0	0
2	0	0	1/4	0	0	1
3	1/2	1	0	0	0	0
4	1/2	0	1/4	1/2	0	0
5	0	0	1/4	0	0	0
6	0	0	1/4	0	1	0

2.3 $M^* = \beta * M + (1 - \beta) * 1/N = 0.8 * M + 1/30$:

	1	2	3	4	5	6
1	1/30	1/30	1/30	13/30	1/30	1/30
2	1/30	1/30	7/30	1/30	1/30	5/6
3	13/30	5/6	1/30	1/30	1/30	1/30
4	13/30	1/30	7/30	13/30	1/30	1/30
5	1/30	1/30	7/30	1/30	1/30	1/30
6	1/30	1/30	7/30	1/30	5/6	1/30

M^* is a stochastic matrix. $r = M^* r$, solve r:
$r = [0.11934435\ 0.19832024\ 0.23972726\ 0.21502754\ 0.08127879\ 0.14630182]^T$
The PageRank scores is r.

2.4 For Personalized-PageRank Node 1:
$M^* = \beta * M + (1 - \beta) * [1\ 0\ 0\ 0\ 0\ 0]^T = 0.8 * M + 0.2 * [1\ 0\ 0\ 0\ 0\ 0]^T$:

	1	2	3	4	5	6
1	0.2	0.2	0.2	0.6	0.2	0.2
2	0	0	0.2	0	0	0.8
3	0.4	0.8	0	0	0	0
4	0.4	0	0.2	0.4	0	0
5	0	0	0.2	0	0	0
6	0	0	0.2	0	0.8	0

M^* is a stochastic matrix. $r = M^* r$, solve r:
$r = [0.30967218\ 0.09916012\ 0.20319697\ 0.27418044\ 0.04063939\ 0.07315091]^T$
The Personalized-PageRankscores for Node 1 is r.