Multidimensional scaling as regression analysis

Okayama University Graduate School

Takashi Shindo

Outline

- An overview of multidimensional scaling (MDS)
- An overview of feature matching model (FMM)
- A constrained MDS as an FMM
    - Feature matching MDS (FM-MDS)
- An FM-MDS as a regression analysis
- Further obstacles

Multidimensional scaling (MDS)

- A method of transforming (dis)similarity data into an arrangement.

- Input is a set of (dis)similarity data between objects.
- Output is coordinates of objects.

- In this presentation, only dissimilarity is considered.

Use of MDS

- The properties of objects can be inferred by searching for meaningful axes in the obtained arrangement.

- Interpretation of the arrangement may be arbitrary.

An example of MDS (input)

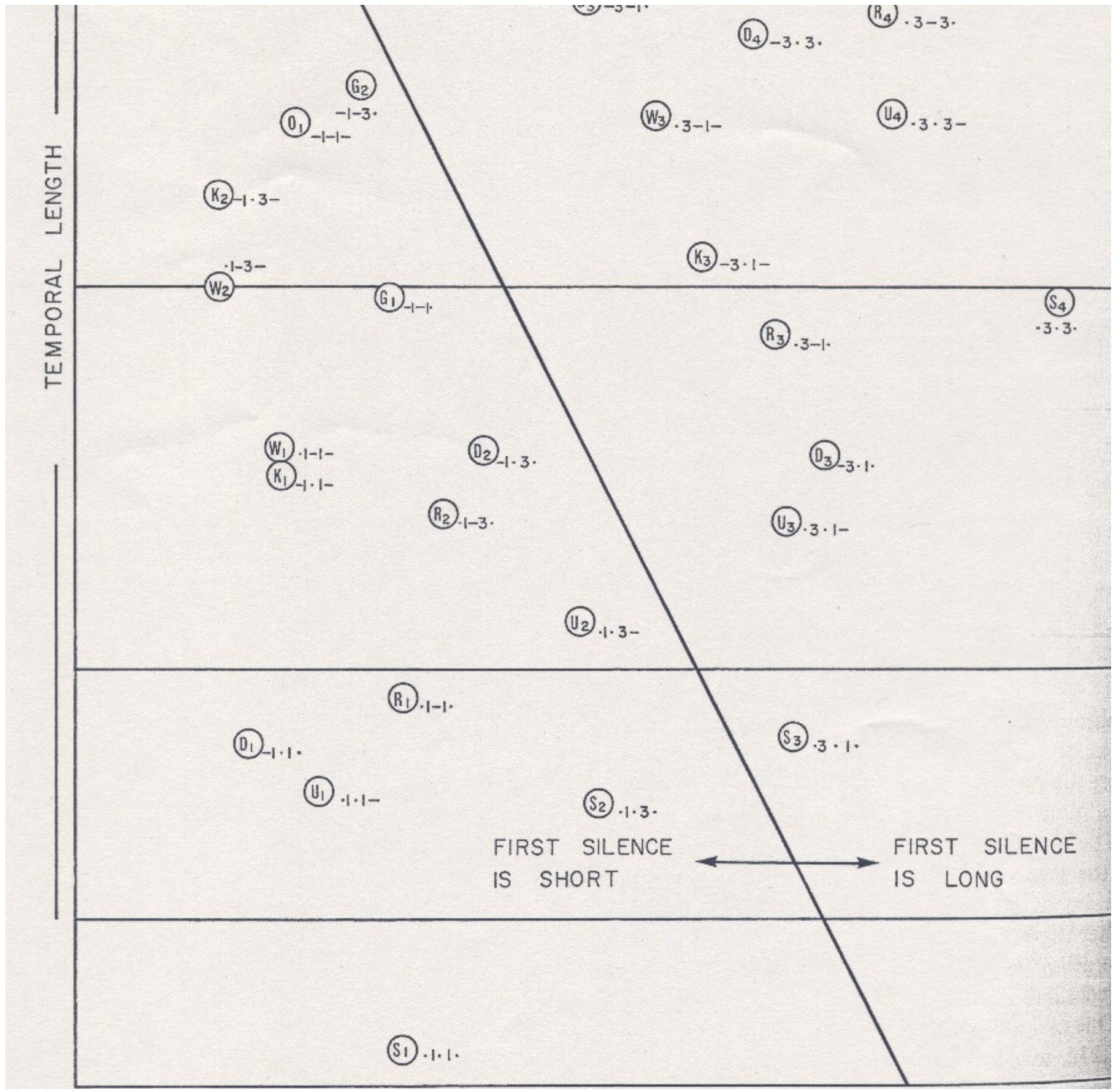
TABLE 2

Percentage of "same" judgments obtained for all ordered pairs of sig

		S_1	U_1	R_1	D_1	W_1	K_1	G_1	O_1	S_2	U_2	R_2	D_2	W_2	K_2	G_2
•|•|•	S_1	97	94	81	66	16	31	06	03	79	18	28	11	01	04	02
•|•|–	U_1	63	98	86	74	91	73	23	16	57	61	49	21	13	09	06
•|–|•	R_1	27	73	94	50	72	62	51	22	42	52	64	49	30	11	16
–|•|•	D_1	47	75	69	96	33	94	67	09	43	13	40	34	11	26	19
•|–|–	W_1	03	40	64	42	94	76	68	70	20	51	73	40	69	49	54
–|•|–	K_1	07	44	33	69	69	95	70	67	19	37	40	56	26	74	27
–|–|•	G_1	01	08	49	50	78	78	93	82	08	14	48	80	23	53	90
–|–|–	O_1	02	10	19	04	31	51	76	97	00	04	12	16	20	50	82
•|•3•	S_2	44	54	62	56	32	41	19	04	94	84	62	54	38	22	07
•|•3–	U_2	03	39	44	15	66	61	29	19	63	94	39	52	77	48	14
•|–3•	R_2	05	16	48	15	59	19	45	28	31	38	95	47	86	36	53
–|•3•	D_2	06	14	28	46	26	51	69	43	47	39	43	94	35	79	73
•|–3–	W_2	02	14	24	09	62	39	27	59	07	54	69	37	94	64	56
–|•3–	K_2	01	03	05	07	22	54	37	56	14	19	23	68	48	98	47
–|–3•	G_2	01	05	10	08	30	28	71	73	08	07	46	57	44	42	93
–|–3–	O_2	00	01	04	03	21	08	35	80	00	04	07	24	55	58	58
•3•|•	S_3	42	38	50	48	23	22	19	06	69	53	31	33	09	03	05
•3•|–	U_3	04	09	30	18	42	52	35	29	30	29	14	21	30	22	05
•3–|•	R_3	05	05	43	08	22	24	38	20	10	17	40	31	28	31	24
–3•|•	D	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Wish, M., (1967). A model for the perception of Morse code-like signals. *Human Factors*, 9, 529-540.

An example of MDS (output)



Wish, M., (1967). A model for the perception of Morse code-like signals. *Human Factors*, 9, 529-540.

Feature matching model (FMM)

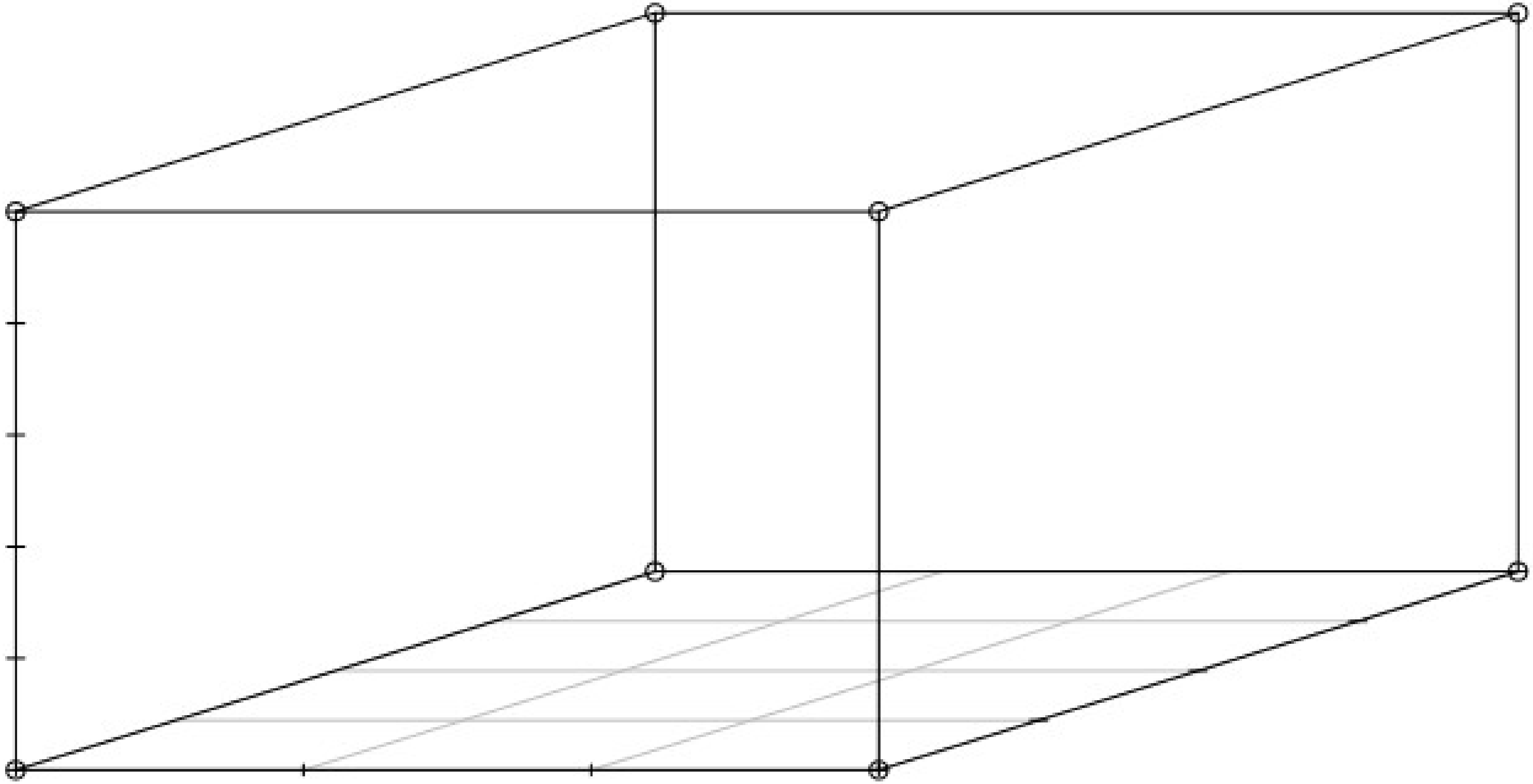
- A regression analysis to explain (dis)similarity with common and distinctive features.
- Interpretation of the result is clear and not arbitrary.
- In this presentation, we consider the case to explain dissimilarity with distinctive features.

Feature matching MDS (FM-MDS)

- We fix the dimension of an arrangement and the order of objects for each dimension.

- By this constraint, each dimension of an arrangement is forced to correspond to a feature.

Idea of FM-MDS



Use of FM-MDS for regression

- We can apply FM-MDS to differences on a scale for regression.
- If we use L1 norm, FM-MDS is equivalent to Hayashi's type I quantification method (Hayashi, 1952).
- A Minkowski norm of higher order is adequate when more contributive features are more dominant to dissimilarity.

Meaning of dominance

- If the maximum norm is used, dissimilarity between a pair of objects is determined by each own feature.
- The idea of sparseness supposes the situation in which a property is determined by a few features.
- Sparseness is assumed to be valid in the field of genetic epidemiology.

Further obstacles

- An efficient solver for FM-MDS is required.

- Risks must be estimated precisely.
  - We need to estimate risks.
  - If sample size << dimension, subjects they have the same set of genotypes will be rare.
  - We cannot use the absolute values of odds ratios.