Journal of Applied Economic Sciences
Volume IV/ Issue 1(7)/ Spring 2009

REPRESENTATION-CONSTRAINED CANONICAL CORRELATION ANALYSIS: A HYBRIDIZATION OF CANONICAL CORRELATION AND PRINCIPAL COMPONENT ANALYSES

Sudhanshu K **MISHRA**
North-Eastern Hill University, Shillong, Meghalaya, **India**
Department of Economics
mishrasknehu@yahoo.com

Abstract:

The classical canonical correlation analysis is extremely greedy to maximize the squared correlation between two sets of variables. As a result, if one of the variables in the dataset-1 is very highly correlated with another variable in the dataset-2, the canonical correlation will be very high irrespective of the correlation among the rest of the variables in the two datasets. We intend here to propose an alternative measure of association between two sets of variables that will not permit the greed of a select few variables in the datasets to prevail upon the fellow variables so much as to deprive the latter of contributing to their representative variables or canonical variates.

Our proposed Representation-Constrained Canonical correlation (RCCCA) Analysis has the Classical Canonical Correlation Analysis (CCCA) at its one end (λ=0) and the Classical Principal Component Analysis (CPCA) at the other (as λ tends to be very large). In between it gives us a compromise solution. By a proper choice of λ, one can avoid hijacking of the representation issue of two datasets by a lone couple of highly correlated variables across those datasets. This advantage of the RCCCA over the CCCA deserves a serious attention by the researchers using statistical tools for data analysis.

Keywords: Representation, constrained, canonical, correlation, principal components, variates, global optimization, particle swarm, ordinal variables, computer program, FORTRAN

JEL Classification: C13, C43, C45, C61, C63, C87

1. Introduction

We begin this paper with reference to a dataset that, when subjected to the classical canonical correlation analysis, gives us the leading (first or largest) canonical correlation which is misleading. It is misleading in the sense that, in this example, the canonical correlation (which is the coefficient of correlation between the two canonical variates, each being a linear weighted combination of the variables in the associated dataset) is, indeed, not a measure of the true association of the variables in the two datasets, but, instead, the datasets have been hijacked by a lone couple of variables across the two datasets.

Table 1.1. Simulated Dataset-1 for Canonical correlation

Sl No.	X_1 or Dataset-1				X_2 or Dataset-2				
	X_{11}	X_{12}	X_{13}	X_{14}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
1	0.7	2.6	0.1	1.7	0.2	0.8	1.6	0.5	1.6
2	1.5	1.7	1.2	1.5	1.6	2.4	2.3	1.4	3.2
3	2.3	0.3	2.7	1.2	2.5	2.9	0.6	1.3	4.8
4	0.6	2.0	0.9	2.8	2.8	2.5	1.1	1.8	1.4
5	0.1	0.9	1.6	1.8	2.2	2.7	2.1	0.2	0.4
6	1.9	1.1	1.7	2.6	1.5	2.2	2.2	2.0	4.0
7	1.0	2.7	2.4	2.7	1.0	0.2	2.0	0.4	2.2
8	1.8	2.9	1.4	0.9	1.7	1.0	1.8	1.2	3.8
9	2.8	0.1	1.8	0.4	2.3	0.6	1.7	0.6	5.8
10	1.4	0.6	2.8	1.4	2.6	1.8	0.8	1.7	3.0
11	1.2	2.5	2.9	0.8	2.1	0.7	1.4	2.3	2.6
12	1.1	1.3	0.2	2.5	0.7	1.5	1.0	2.2	2.4
13	3.0	1.9	1.1	1.6	0.1	0.1	2.7	3.0	6.2

Sl No.	X_1 or Dataset-1				X_2 or Dataset-2				
	X_{11}	X_{12}	X_{13}	X_{14}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
14	2.0	0.8	0.6	1.3	1.9	0.5	0.4	0.8	4.2
15	1.6	2.2	2.6	1.9	1.4	1.3	1.3	2.5	3.4
16	2.9	0.7	1.9	2.9	2.4	1.2	2.5	2.1	6.0
17	1.3	1.4	2.0	0.2	1.8	2.8	0.3	2.6	2.8
18	0.8	0.2	2.3	2.0	2.9	1.4	3.0	0.7	1.8
19	1.7	0.5	1.3	0.1	2.0	0.9	2.9	1.5	3.6
20	2.1	2.4	0.7	0.5	0.9	2.3	0.7	0.3	4.4
21	2.5	1.0	3.0	2.2	1.2	2.6	2.6	1.0	5.2
22	2.2	2.8	2.5	0.7	3.0	3.0	0.2	1.9	4.6
23	0.5	0.4	0.8	1.0	0.8	0.4	0.1	1.1	1.2
24	2.7	2.1	1.5	2.3	1.1	1.1	0.9	2.7	5.6
25	2.4	1.8	0.5	0.3	2.7	1.6	2.8	0.1	5.0
26	0.2	1.6	0.3	1.1	0.6	0.3	2.4	2.8	0.6
27	0.9	2.3	0.4	0.6	1.3	1.7	1.5	2.4	2.0
28	2.6	3.0	2.2	3.0	0.5	1.9	1.9	1.6	5.4
29	0.4	1.2	1.0	2.4	0.4	2.0	0.5	2.9	1.0
30	0.3	1.5	2.1	2.1	0.3	2.1	1.2	0.9	0.8

In Table 1.1 the dataset X is presented which is a pooled set of two datasets, X_1 and X_2, such that $X=[X_1|X_2]$. The first dataset has m_1 (=4) variables and the second dataset has m_2 (=5) variables, each in n (=30) observations. These seemingly normal datasets, when subjected to the classical canonical correlation analysis, yield canonical correlation between the composite variables, z_1 and z_2 (the canonical variates), $r(z_1,z_2)=1.0$: $z_1=\sum_{j=1}^{4} w_j x_{1j}; x_{ij}\in X_1$; $z_2=\sum_{j=1}^{5} w_j x_{2j}; x_{2j}\in X_2$. The weight vectors are: w_1=(1, 0, 0, 0, 0) and w_2=(0, 0, 0, 0, 1). This anomalous situation has arisen due to the fact that x_{25} is perfectly linearly dependent on x_{11} and the canonical correlation, $r(z_1,z_2)$, is in fact $r(x_{11},x_{25})$. Other variables have no contribution to z_1 or z_2. It follows, therefore, that z_1 and z_2 do not represent other variables in X_1 and X_2. Nor is the canonical correlation, $r(z_1,z_2)$, a correlation between the two sets, X_1 and X_2, in any relevant or significant sense. Thus, the leading canonical correlation may deceive us if we are only a little less careful to look into the correlation matrix encompassing all variables.

Such examples may be multiplied *ad infinitum*. If one is cautious, the anomalous cases can be detected. However, such cases, if not detected, make scientific analysis and interpretation of empirical results rather hazardous. One may easily be misled to a conclusion that such two datasets are highly correlated while the truth may be quite far from it.

2. Objectives of the Present Work

We intend here to propose an alternative measure of association between two sets of variables that will not permit the greed of a select few variables in the datasets to prevail upon the fellow variables so much as to deprive the latter of contributing their say and share to the representative variables (ς_1 and ς_2), which they make by their participation in the linear combination. We may not call $\varsigma_1=\sum_{j=1}^{m_1}\omega_{1j}x_{1j}$ and $\varsigma_2=\sum_{j=1}^{m_2}\omega_{2j}x_{2j}$ the canonical variables (defined before as $z_1=\sum_{j=1}^{4} w_j x_{1j}$; $z_2=\sum_{j=1}^{5} w_j x_{2j}$ obtained from the classical canonical correlation analysis).

In the classical canonical correlation analysis the objective is to maximize $r^2(z_1,z_2): z_1=\sum_{j=1}^{m_1} w_{1j}x_{1j}; z_2=\sum_{j=1}^{m_2} w_{2j}x_{2j}$ irrespective of $r(z_1,x_{1j}): x_{1j}\in X_1$ and $r(z_2,x_{2j}): x_{2j}\in X_2$, and, therefore, $r^2(z_1,z_2)$ is subject to an unconstrained maximization. However, in the method that we are proposing here, the objective will be to maximize $r^2(\varsigma_1,\varsigma_2)$: $\varsigma_1=\sum_{j=1}^{m_1}\omega_{1j}x_{1j}$ and $\varsigma_2=\sum_{j=1}^{m_2}\omega_{2j}x_{2j}$ with certain constraints in terms of $r(\varsigma_1,x_{1j}): x_{1j}\in X_1$ and $r(\varsigma_2,x_{2j}): x_{2j}\in X_2$. These constraints would ensure the representativeness of ς_1 to X_1 and that of ς_2 to X_2. Hence, the proposed method may be called the *Representation-Constrained Canonical Correlation Analysis*.

3. The Nature and Implications of the Proposed Constraints

There are a number of ways in which the canonical variates can be constrained insofar as their association and concordance with their fellow variables in their respective native datasets are concerned. In other words, their representativeness to their native datasets can be defined variously. We discuss here some of the alternatives in terms of correlation as a measure of representativeness.

(i) Mean absolute correlation principle: A (constrained) canonical variate $\varsigma_a = \sum_{j=1}^{m_a} \omega_{aj} x_{aj}; x_{aj} \in X_a$ is a better representative of X_a if the mean absolute correlation, $\sum_{j=1}^{m_a} |r(\varsigma_a, x_{aj})|$, is larger. This approach is equalitarian in effect.

(ii) Mean squared correlation principle: A (constrained) canonical variate $\varsigma_a = \sum_{j=1}^{m_a} \omega_{aj} x_{aj}; x_{aj} \in X_a$ is a better representative of X_a if the mean squared correlation, $\sum_{j=1}^{m_a} r^2(\varsigma_a, x_{aj})$, is larger. This approach is elitist in effect, favouring dominant members.

(iii) Minimal absolute correlation principle: A (constrained) canonical variate $\varsigma_a = \sum_{j=1}^{m_a} \omega_{aj} x_{aj}; x_{aj} \in X_a$ is a better representative of X_a if the minimal absolute correlation, $\min_j [|r(\varsigma_a, x_{aj})|]$, is larger. A larger $\min_j [|r(\varsigma_a, x_{aj})|]$ implies that the minimal squared correlation, $\min_j [r^2(\varsigma_a, x_{aj})]$, is larger. This approach is in favour of the weak.

These three approaches lead to three alternative objective functions:

(i). Maximize $r^2(\varsigma_1,\varsigma_2) + \lambda[\sum_{j=1}^{m_1} |r(\varsigma_1, x_{1j})| / m_1 + \sum_{j=1}^{m_2} |r(\varsigma_2, x_{2j})| / m_2] : \varsigma_1 = \sum_{j=1}^{m_1} \omega_{1j} x_{1j};\ \varsigma_2 = \sum_{j=1}^{m_2} \omega_{2j} x_{2j}.$

(ii). Maximize $r^2(\varsigma_1,\varsigma_2) + \lambda[\sum_{j=1}^{m_1} r^2(\varsigma_1, x_{1j}) / m_1 + \sum_{j=1}^{m_2} r^2(\varsigma_2, x_{2j}) / m_2] : \varsigma_1 = \sum_{j=1}^{m_1} \omega_{1j} x_{1j};\ \varsigma_2 = \sum_{j=1}^{m_2} \omega_{2j} x_{2j}.$

(iii). Maximize $r^2(\varsigma_1,\varsigma_2) + \lambda\left[\min_j [|r(\varsigma_1, x_{1j})|] + \min_j [|r(\varsigma_2, x_{2j})|]\right] : \varsigma_1 = \sum_{j=1}^{m_1} \omega_{1j} x_{1j};\ \varsigma_2 = \sum_{j=1}^{m_2} \omega_{2j} x_{2j}.$

In these objective functions, the value of λ may be chosen subjectively. If $\lambda = 0$, the objective function would degenerate to the classical canonical correlation analysis, but λ has no upper bound. Also note that if the first term is $|r(\varsigma_1,\varsigma_2)|$ rather than $r^2(\varsigma_1,\varsigma_2)$ and $\lambda \neq 0$, its implied weight vis-à-vis the second term increases since $|r(\varsigma_1,\varsigma_2)| > r^2(\varsigma_1,\varsigma_2)$ for $|r| < 1$.

4. The Method of Optimization

The classical canonical correlation analysis [Hotelling, (1936)] sets up the objective function to maximize $r^2(\varsigma_1,\varsigma_2) : \varsigma_1 = \sum_{j=1}^{m_1} \omega_{1j} x_{1j};\ \varsigma_2 = \sum_{j=1}^{m_2} \omega_{2j} x_{2j}$ and using the calculus methods of maximization resolves the problem to finding out the largest eigenvalue and the associated eigenvector of the matrix, $[X_1'X_1]^{-1} X_1'X_2 [X_2'X_2]^{-1} X_2'X_1$. The largest eighen-value turns out to be the leading $r^2(z_1, z_2) : z_1 = \sum_{j=1}^{m_1} w_{1j} x_{1j};\ z_2 = \sum_{j=1}^{m_2} w_{2j} x_{2j}$, and the standardized eigenvector is used to obtain w_1 and w_2. However, a general calculus-based method cannot be applied to maximize the (arbitrary) objective function set up for the constrained canonical correlation analysis. At any rate, the first and the third objective functions are not amenable to maximization by the calculus-based methods.

We choose, therefore, to use a relatively new and more versatile method of (global) optimization, namely, the Particle Swarm Optimization (PSO) proposed by Eberhart and Kennedy (1995). A lucid description of its foundations is available in Fleischer (2005). The PSO is a biologically inspired population-based stochastic search method modeled on the ornithological observations, simulating the behavior of members of the flocks of birds in searching food and communicating among themselves. It is in conformity with the principles of decentralized decision making [Hayek, (1948); (1952)] leading to self-organization and macroscopic order. The effectiveness of PSO has been very encouraging in solving extremely difficult and varied types of nonlinear optimization problems [Mishra, (2006)]. We have used a particular variant of the PSO called the Repulsive Particle Swarm Optimization [Urfalioglu, (2004)].

5. Findings and Discussion

We have subjected the data in Table 1.1 to the representation-constrained canonical correlation analysis with the three alternative objective functions elaborated in section-III. The first term, measuring the degree of association between the two datasets, X_1 and X_2, is in the squared form, that is $r^2(\varsigma_1,\varsigma_2)$, although we have reported its positive square root (=$|r(\varsigma_1,\varsigma_2)|$) in Table 1.2. The three objective functions have been optimized for the different values of λ, varying from zero to 50 with an increment of 0.5. For the first objective function, the values of $|r(\varsigma_1,\varsigma_2)|$, mean absolute $r(\varsigma_1,x_1)$ and mean absolute $r(\varsigma_2,x_2)$ at different values of λ have been plotted in Figure 1.1. Similarly, for the second objective function, the values of $|r(\varsigma_1,\varsigma_2)|$, mean squared $r(\varsigma_1,x_1)$ and mean squared $r(\varsigma_2,x_2)$ at different values of λ have been plotted in Figure1.2., Figure 1.3 presents $|r(\varsigma_1,\varsigma_2)|$, minimum absolute $r(\varsigma_1,x_1)$ and minimum absolute $r(\varsigma_2,x_2)$ relating to the 3rd objective maximized at different values of λ.

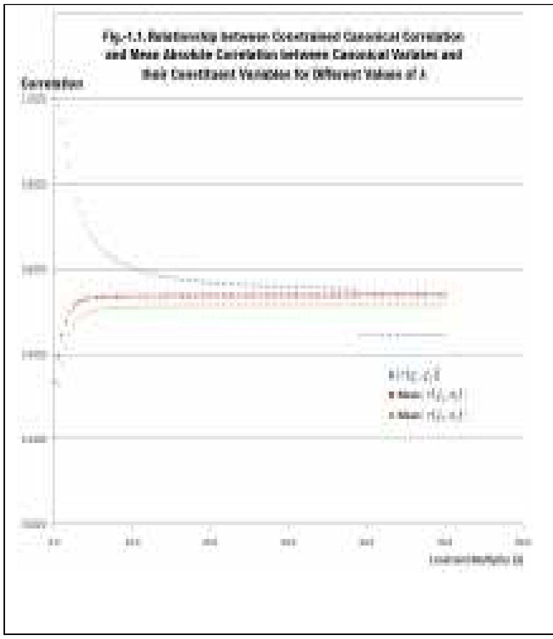
Table 1.2. Relationship between Constrained Canonical Correlation and Representation Correlation between Canonical Variates and their Constituent Variables for Different Values of λ

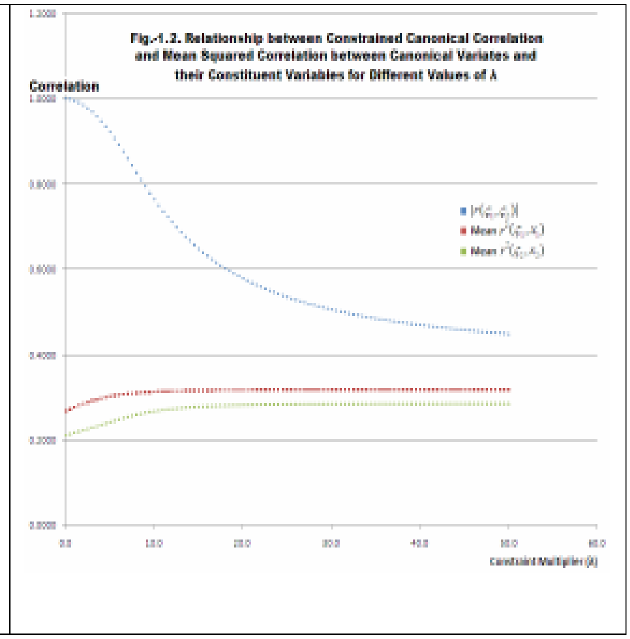
Sl No	λ	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Absolute $r(\varsigma_1,x_1)$	Mean Absolute $r(\varsigma_2,x_2)$	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Squared $r(\varsigma_1,x_1)$	Mean Squared $r(\varsigma_2,x_2)$	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Minimum Absolute $r(\varsigma_1,x_1)$	Minimum Absolute $r(\varsigma_2,x_2)$
1	0.0	1.0000	0.3342	0.2814	1.0000	0.2668	0.2121	1.0000	0.0234	0.0246
2	0.5	0.9831	0.3942	0.3254	0.9990	0.2717	0.2152	0.9755	0.1615	0.1361
3	1.0	0.9440	0.4434	0.3785	0.9961	0.2763	0.2183	0.8223	0.3921	0.2671
4	1.5	0.8942	0.4772	0.4188	0.9916	0.2805	0.2214	0.5072	0.5302	0.4618
5	2.0	0.8432	0.4992	0.4479	0.9855	0.2843	0.2244	0.4662	0.5319	0.4853
6	2.5	0.7975	0.5128	0.4679	0.9780	0.2878	0.2275	0.4556	0.5337	0.4889
7	3.0	0.7597	0.5210	0.4813	0.9691	0.2909	0.2306	0.4473	0.5338	0.4917
8	3.5	0.7298	0.5259	0.4902	0.9590	0.2938	0.2338	0.4296	0.5337	0.4968
9	4.0	0.7060	0.5290	0.4962	0.9477	0.2964	0.2369	0.4349	0.5334	0.4954
10	4.5	0.6870	0.5310	0.5005	0.9352	0.2987	0.2401	0.4230	0.5335	0.4978
11	5.0	0.6715	0.5323	0.5036	0.9217	0.3008	0.2433	0.4342	0.5337	0.4955
12	5.5	0.6590	0.5333	0.5058	0.9073	0.3027	0.2464	0.4359	0.5338	0.4950
13	6.0	0.6483	0.5339	0.5076	0.8921	0.3044	0.2495	0.4404	0.5338	0.4940
14	6.5	0.6394	0.5345	0.5089	0.8762	0.3059	0.2525	0.3743	0.5389	0.4963
15	7.0	0.6318	0.5348	0.5100	0.8599	0.3072	0.2554	0.4170	0.5337	0.4994
16	7.5	0.6251	0.5351	0.5108	0.8434	0.3083	0.2581	0.4175	0.5338	0.4992
17	8.0	0.6193	0.5354	0.5115	0.8270	0.3094	0.2607	0.4278	0.5338	0.4970
18	8.5	0.6142	0.5356	0.5121	0.8106	0.3102	0.2630	0.4167	0.5335	0.4990
19	9.0	0.6098	0.5357	0.5126	0.7945	0.3110	0.2652	0.4293	0.5337	0.4967
20	9.5	0.6056	0.5358	0.5130	0.7789	0.3117	0.2672	0.4206	0.5339	0.4986
21	10.0	0.6019	0.5360	0.5133	0.7641	0.3122	0.2690	0.3746	0.5389	0.4962
22	10.5	0.5988	0.5360	0.5136	0.7495	0.3127	0.2706	0.2904	0.5023	0.4748
23	11.0	0.5958	0.5361	0.5139	0.7359	0.3132	0.2721	0.4167	0.5338	0.4990
24	11.5	0.5931	0.5362	0.5141	0.7227	0.3136	0.2734	0.4201	0.4789	0.4281
25	12.0	0.5906	0.5362	0.5143	0.7103	0.3139	0.2746	0.4206	0.5338	0.4987
26	12.5	0.5884	0.5363	0.5144	0.6985	0.3142	0.2756	0.5150	0.4781	0.3664
27	13.0	0.5861	0.5363	0.5146	0.6872	0.3145	0.2766	0.4167	0.5337	0.4993
28	13.5	0.5842	0.5364	0.5147	0.6764	0.3148	0.2774	0.3745	0.5389	0.4964
29	14.0	0.5826	0.5364	0.5148	0.6665	0.3150	0.2782	0.3742	0.5390	0.4963
30	14.5	0.5807	0.5364	0.5150	0.6570	0.3152	0.2789	0.4022	0.4648	0.4532
31	15.0	0.5791	0.5365	0.5151	0.6478	0.3154	0.2795	0.4170	0.5338	0.4991
32	15.5	0.5778	0.5365	0.5151	0.6390	0.3155	0.2801	0.4179	0.5003	0.4860
33	16.0	0.5765	0.5365	0.5152	0.6310	0.3157	0.2806	0.2791	0.5387	0.4990
34	16.5	0.5751	0.5365	0.5153	0.6231	0.3158	0.2810	0.3992	0.4764	0.4347
35	17.0	0.5739	0.5365	0.5154	0.6158	0.3159	0.2815	0.3742	0.5388	0.4964
36	17.5	0.5728	0.5366	0.5154	0.6088	0.3160	0.2819	0.0285	0.4457	0.4501
37	18.0	0.5715	0.5366	0.5155	0.6021	0.3161	0.2822	0.2794	0.5389	0.4992
38	18.5	0.5706	0.5366	0.5155	0.5960	0.3162	0.2825	0.3811	0.4744	0.4599
39	19.0	0.5697	0.5366	0.5156	0.5898	0.3163	0.2828	0.3741	0.5389	0.4963
40	19.5	0.5688	0.5366	0.5156	0.5840	0.3164	0.2831	0.3743	0.5389	0.4962
41	20.0	0.5680	0.5366	0.5157	0.5783	0.3165	0.2834	0.3345	0.4838	0.3983
42	20.5	0.5671	0.5366	0.5157	0.5732	0.3166	0.2836	0.2795	0.5389	0.4994
43	21.0	0.5663	0.5366	0.5157	0.5682	0.3166	0.2838	0.4194	0.4718	0.4439
44	21.5	0.5655	0.5366	0.5158	0.5632	0.3167	0.2840	0.3746	0.5389	0.4963
45	22.0	0.5650	0.5367	0.5158	0.5587	0.3167	0.2842	0.5496	0.5103	0.3823
46	22.5	0.5643	0.5367	0.5158	0.5542	0.3168	0.2843	0.2539	0.5138	0.4743
47	23.0	0.5635	0.5367	0.5158	0.5499	0.3168	0.2845	0.2795	0.5390	0.4993
48	23.5	0.5630	0.5367	0.5159	0.5459	0.3169	0.2846	0.2865	0.4643	0.4394

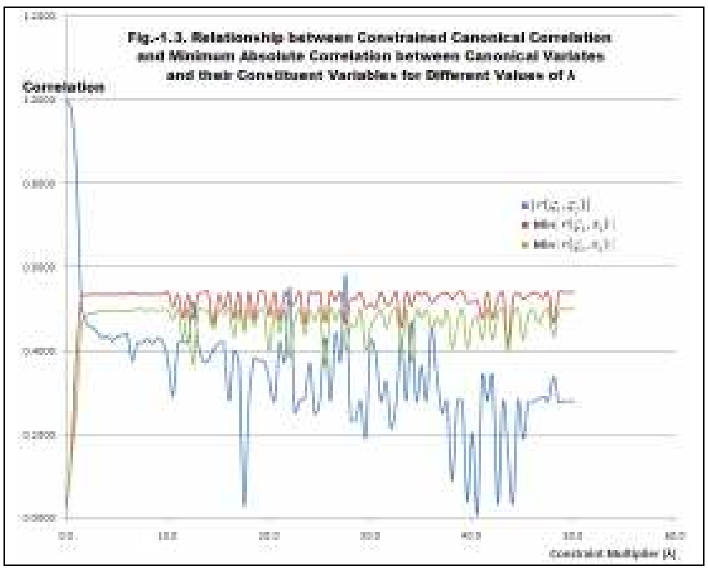
49	24.0	0.5623	0.5367	0.5159	0.5419	0.3169	0.2848	0.3688	0.5389	0.4944
50	24.5	0.5618	0.5367	0.5159	0.5383	0.3170	0.2849	0.2490	0.5347	0.4720
51	25.0	0.5612	0.5367	0.5159	0.5347	0.3170	0.2850	0.2792	0.5387	0.4994
52	25.5	0.5607	0.5367	0.5159	0.5312	0.3170	0.2851	0.4305	0.4684	0.3653
53	26.0	0.5603	0.5367	0.5160	0.5280	0.3171	0.2852	0.2793	0.5387	0.4993
54	26.5	0.5597	0.5367	0.5160	0.5249	0.3171	0.2853	0.4418	0.5176	0.4731
55	27.0	0.5592	0.5367	0.5160	0.5219	0.3171	0.2854	0.3741	0.5388	0.4963
56	27.5	0.5589	0.5367	0.5160	0.5186	0.3171	0.2855	0.5795	0.4661	0.4031
57	28.0	0.5584	0.5367	0.5160	0.5160	0.3172	0.2856	0.2335	0.5213	0.4604
58	28.5	0.5581	0.5367	0.5160	0.5131	0.3172	0.2857	0.2335	0.5213	0.4604
59	29.0	0.5575	0.5367	0.5161	0.5103	0.3172	0.2858	0.2790	0.5388	0.4993
60	29.5	0.5572	0.5367	0.5161	0.5080	0.3172	0.2858	0.1922	0.5023	0.4015
61	30.0	0.5568	0.5367	0.5161	0.5054	0.3173	0.2859	0.4223	0.5119	0.4564
62	30.5	0.5564	0.5367	0.5161	0.5030	0.3173	0.2859	0.3929	0.5016	0.4801
63	31.0	0.5561	0.5367	0.5161	0.5008	0.3173	0.2860	0.2795	0.5390	0.4993
64	31.5	0.5558	0.5367	0.5161	0.4987	0.3173	0.2861	0.3260	0.5081	0.4567
65	32.0	0.5555	0.5367	0.5161	0.4964	0.3173	0.2861	0.2140	0.5156	0.4897
66	32.5	0.5549	0.5367	0.5161	0.4942	0.3173	0.2862	0.2793	0.5389	0.4992
67	33.0	0.5547	0.5367	0.5161	0.4921	0.3174	0.2862	0.4277	0.4566	0.4137
68	33.5	0.5545	0.5367	0.5161	0.4902	0.3174	0.2863	0.2794	0.5389	0.4993
69	34.0	0.5542	0.5367	0.5161	0.4883	0.3174	0.2863	0.4708	0.5056	0.3723
70	34.5	0.5539	0.5367	0.5162	0.4865	0.3174	0.2863	0.2787	0.5388	0.4988
71	35.0	0.5539	0.5367	0.5162	0.4846	0.3174	0.2864	0.3639	0.5312	0.4787
72	35.5	0.5534	0.5367	0.5162	0.4830	0.3174	0.2864	0.2793	0.5389	0.4992
73	36.0	0.5532	0.5367	0.5162	0.4814	0.3174	0.2864	0.4560	0.5133	0.4533
74	36.5	0.5528	0.5367	0.5162	0.4796	0.3174	0.2865	0.3375	0.5282	0.4788
75	37.0	0.5524	0.5368	0.5162	0.4780	0.3174	0.2865	0.2504	0.5345	0.4600
76	37.5	0.5524	0.5368	0.5162	0.4765	0.3175	0.2865	0.2784	0.5380	0.4988
77	38.0	0.5521	0.5368	0.5162	0.4749	0.3175	0.2866	0.0886	0.5222	0.4078
78	38.5	0.5520	0.5368	0.5162	0.4733	0.3175	0.2866	0.2791	0.5372	0.4631
79	39.0	0.4469	0.5394	0.5163	0.4721	0.3175	0.2866	0.2795	0.5389	0.4992
80	39.5	0.4468	0.5394	0.5163	0.4707	0.3175	0.2866	0.0385	0.5148	0.4071
81	40.0	0.4467	0.5394	0.5163	0.4693	0.3175	0.2867	0.2028	0.5160	0.4721
82	40.5	0.4463	0.5394	0.5163	0.4681	0.3175	0.2867	0.0080	0.5182	0.4812
83	41.0	0.4463	0.5394	0.5163	0.4666	0.3175	0.2867	0.3389	0.4771	0.4282
84	41.5	0.4461	0.5394	0.5163	0.4653	0.3175	0.2867	0.2795	0.5389	0.4994
85	42.0	0.4460	0.5394	0.5163	0.4644	0.3175	0.2868	0.3389	0.4771	0.4282
86	42.5	0.4458	0.5394	0.5163	0.4631	0.3175	0.2868	0.0338	0.5248	0.4897
87	43.0	0.4456	0.5394	0.5163	0.4617	0.3175	0.2868	0.2793	0.5389	0.4993
88	43.5	0.4454	0.5394	0.5163	0.4606	0.3175	0.2868	0.1597	0.4139	0.3977
89	44.0	0.4453	0.5394	0.5163	0.4593	0.3176	0.2868	0.0338	0.5248	0.4897
90	44.5	0.4452	0.5394	0.5163	0.4586	0.3176	0.2869	0.2794	0.5389	0.4994
91	45.0	0.4451	0.5394	0.5163	0.4576	0.3176	0.2869	0.1880	0.5229	0.4274
92	45.5	0.4450	0.5394	0.5163	0.4564	0.3176	0.2869	0.2733	0.5300	0.4848
93	46.0	0.4448	0.5394	0.5163	0.4555	0.3176	0.2869	0.2786	0.5389	0.4991
94	46.5	0.4447	0.5394	0.5163	0.4547	0.3176	0.2869	0.2822	0.5354	0.4665
95	47.0	0.4445	0.5394	0.5163	0.4535	0.3176	0.2869	0.2898	0.5252	0.4905
96	47.5	0.4444	0.5394	0.5163	0.4527	0.3176	0.2869	0.2796	0.5389	0.4993
97	48.0	0.4444	0.5394	0.5163	0.4510	0.3176	0.2870	0.3372	0.4676	0.4344
98	48.5	0.4442	0.5394	0.5163	0.4509	0.3176	0.2870	0.2768	0.5389	0.4985
99	49.0	0.4440	0.5394	0.5163	0.4500	0.3176	0.2870	0.2792	0.5388	0.4993
100	49.5	0.4439	0.5394	0.5163	0.4491	0.3176	0.2870	0.2790	0.5389	0.4993
101	50.0	0.4438	0.5394	0.5163	0.4480	0.3176	0.2870	0.2784	0.5390	0.4989

From Figure 1.1 and Figure 1.2 it is clear that for increasing values of λ, the value of $|r(\varsigma_1,\varsigma_2)|$ decreases monotonically, while the values of mean absolute (or squared) $r(\varsigma_1,x_1)$ and mean absolute (or squared) $r(\varsigma_2,x_2)$ increase monotonically. All of them exhibit asymptotic tendencies. However, for the third objective function the monotonicity of all the correlation functions is lost (shown in Figure 1.3). Of course, the trends in minimum absolute $r(\varsigma_1,x_1)$ and minimum absolute $r(\varsigma_2,x_2)$ are clearly observable. These observations may be useful to the choice of λ. For the case that we are presently dealing with, the value of λ need not exceed 10 to assure a fairly satisfactory representation of the two datasets by the corresponding canonical variates.

In particular, optimization of the second objective function has shown that the values of mean squared $r(\varsigma_1,x_1)$ and mean squared $r(\varsigma_2,x_2)$ exhibit asymptotic tendencies. For λ=50, the mean squared $r(\varsigma_1,x_1)$ is 0.3176 while the mean squared $r(\varsigma_2,x_2)$ is 0.2870.







Now, let us digress for a while to compute the first principal components of X_1 and X_2 (from the data given in Table 1.1). We find that for X_1 the sum of squared correlation (component loadings) of the component score (ξ_1) with its constituent variables is 0.317757. In other words, the first eigenvalue of the inter-correlation matrix R_1 obtained from X_1 is 1.271029, which divided by 4 (order of R_1) gives 0.317757. This is, in a way, a measure of representation of X_1 by its first principal component. Similarly, for X_2 the sum of squared correlation of the component score (ξ_2) with its constituent variables is 0.287521.

We resume our discussion for comparing these results (obtained from the Principal Component Analysis) with the results of our proposed representation-constrained canonical correlation analysis. We observe that the asymptotic tendencies of mean squared $r(\varsigma_1,x_1)$ and mean squared $r(\varsigma_2,x_2)$ clearly point to the explanatory powers of the first principal components of X_1 and X_2 respectively.

However, if we compute the coefficient of correlation between the two component scores ($r(\xi_1,\xi_2)$=0.390767) and compare it with the constrained canonical correlation ($r(\varsigma_1,\varsigma_2)$=0.4480 for λ=50) we find that the latter is larger. Then, is the constrained canonical correlation analysis a hybrid of the classical canonical correlation and principal component analyses which has better properties of representation of data than its parents?

We conduct another experiment with the dataset presented in Table 2.1. We find that ξ_1 for X_1 has the representation power 0.333261 (eigenvalue=1.333042) while ξ_2 for X_2 has the representation power 0.382825 (eigenvalue=1.914123). The $r(\xi_1,\xi_2)$ is 0.466513. On the other hand, results of the constrained canonical correlation (for λ=49) are: mean squared $r(\varsigma_1,x_1)$= 0.33317; mean squared $r(\varsigma_2,x_2)$=0.38270 and the representation-constrained canonical correlation, $r(\varsigma_1,\varsigma_2)$= 0.48761. These findings are corroborative to our earlier results with regard to the dataset in Table 1.1.

Table 2.1. Simulated Dataset-2 for Canonical correlation

Sl	X_1 or Dataset-1				X_2 or Dataset-2					Sl	X_1 or Dataset-1				X_2 or Dataset-2				
No.	X_{11}	X_{12}	X_{13}	X_{14}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	No.	X_{11}	X_{12}	X_{13}	X_{14}	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
1	2.7	1.9	2.4	1.2	2.6	2.3	1.5	0.1	6.6	16	1.4	1.4	0.2	0.4	1.1	2.2	2.2	2.6	2.4
2	1.7	0.1	0.8	1.8	0.4	0.2	2.3	0.2	0.1	17	1.5	0.4	2.2	1.9	1.9	0.6	2.1	1.9	5.5
3	0.2	2.8	2.6	0.9	1.3	2.0	2.0	0.3	7.3	18	0.6	1.3	2.5	2.8	2.8	2.5	2.7	2.2	4.9
4	0.4	0.3	1.1	0.2	1.5	1.3	1.1	1.8	3.3	19	2.5	1.1	0.1	1.1	2.5	1.0	1.0	2.9	3.8
5	0.9	1.8	1.6	1.4	0.8	1.2	2.4	2.4	5.7	20	1.0	2.3	1.8	1.5	2.9	1.8	1.6	2.0	5.8
6	0.5	0.9	2.7	0.7	1.4	1.6	1.2	3.0	6.2	21	0.8	1.7	1.0	1.6	1.6	2.4	0.6	1.4	4.5
7	2.0	1.0	2.9	1.7	0.3	0.1	0.4	1.1	5.4	22	0.3	1.2	2.1	0.3	2.0	1.9	0.7	0.9	4.5
8	0.1	1.6	0.5	2.7	0.7	2.1	1.3	1.7	3.1	23	1.3	0.7	1.3	2.4	2.2	0.7	0.8	1.0	3.4
9	1.2	0.6	2.8	1.0	0.1	0.9	0.1	0.8	3.7	24	2.6	1.5	2.3	0.6	1.7	2.9	2.9	2.5	7.3
10	2.9	2.1	0.4	0.8	0.5	0.3	1.7	0.4	4.7	25	3.0	2.6	1.2	3.0	2.7	2.6	2.8	1.5	7.2
11	0.7	0.5	0.6	1.3	2.1	0.5	0.3	0.7	0.7	26	1.1	2.2	0.7	2.5	2.4	0.8	2.6	1.2	3.8
12	2.8	2.5	1.5	2.9	2.3	2.8	3.0	1.6	6.5	27	1.8	2.0	1.9	2.2	1.8	1.7	1.8	0.6	6.1
13	2.2	0.2	1.7	2.3	3.0	1.1	0.5	2.7	3.9	28	1.9	2.7	3.0	2.0	1.0	1.4	1.4	0.5	9.5
14	2.1	0.8	0.9	2.6	0.9	2.7	2.5	2.1	3.6	29	1.6	2.4	0.3	0.5	0.2	0.4	0.2	2.3	6.5
15	2.3	3.0	1.4	0.1	0.6	3.0	0.9	2.8	8.4	30	2.4	2.9	2.0	2.1	1.2	1.5	1.9	1.3	7.6

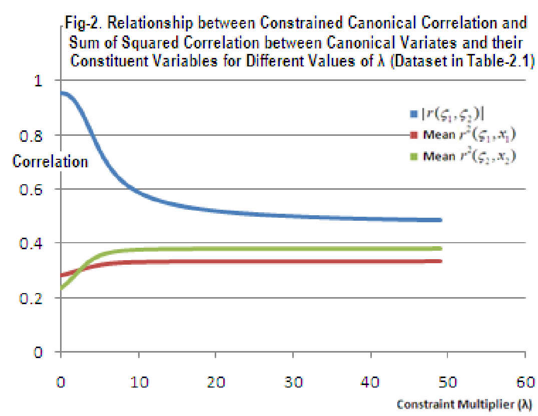


Table 2.2. Relationship between Constrained Canonical Correlation and Representation Correlation between Canonical Variates and their Constituent Variables for Different Values of λ (Dataset in Table-2.1)

Sl No.	λ	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Squared $r(\varsigma_1,x_1)$	Mean Squared $r(\varsigma_2,x_2)$	Sl No.	λ	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Squared $r(\varsigma_1,x_1)$	Mean Squared $r(\varsigma_2,x_2)$
1	0.0	0.95772	0.28514	0.23519	26	25.0	0.50983	0.33290	0.38230
2	1.0	0.94904	0.29212	0.26011	27	26.0	0.50804	0.33293	0.38234
3	2.0	0.91881	0.29987	0.28983	28	27.0	0.50638	0.33296	0.38238
4	3.0	0.86701	0.30782	0.31901	29	28.0	0.50475	0.33298	0.38241
5	4.0	0.80425	0.31506	0.34197	30	29.0	0.50340	0.33300	0.38244
6	5.0	0.74448	0.32066	0.35709	31	30.0	0.50203	0.33302	0.38247
7	6.0	0.69541	0.32452	0.36618	32	31.0	0.50086	0.33304	0.38249
8	7.0	0.65777	0.32703	0.37155	33	32.0	0.49966	0.33305	0.38252
9	8.0	0.62930	0.32867	0.37482	34	33.0	0.49858	0.33306	0.38254
10	9.0	0.60764	0.32976	0.37690	35	34.0	0.49755	0.33308	0.38256
11	10.0	0.59071	0.33052	0.37828	36	35.0	0.49659	0.33309	0.38257
12	11.0	0.57730	0.33106	0.37924	37	36.0	0.49571	0.33310	0.38259

Sl No.	λ	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Squared $r(\varsigma_1,x_1)$	Mean Squared $r(\varsigma_2,x_2)$	Sl No.	λ	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Squared $r(\varsigma_1,x_1)$	Mean Squared $r(\varsigma_2,x_2)$
13	12.0	0.56634	0.33146	0.37993	38	37.0	0.49492	0.33311	0.38260
14	13.0	0.55733	0.33176	0.38044	39	38.0	0.49409	0.33311	0.38261
15	14.0	0.54983	0.33199	0.38083	40	39.0	0.49333	0.33312	0.38262
16	15.0	0.54338	0.33217	0.38113	41	40.0	0.49265	0.33313	0.38263
17	16.0	0.53786	0.33232	0.38137	42	41.0	0.49193	0.33314	0.38264
18	17.0	0.53310	0.33244	0.38156	43	42.0	0.49132	0.33314	0.38265
19	18.0	0.52896	0.33253	0.38171	44	43.0	0.49074	0.33315	0.38266
20	19.0	0.52523	0.33262	0.38185	45	44.0	0.49018	0.33315	0.38267
21	20.0	0.52196	0.33268	0.38195	46	45.0	0.48958	0.33316	0.38268
22	21.0	0.51904	0.33274	0.38205	47	46.0	0.48912	0.33316	0.38268
23	22.0	0.51638	0.33279	0.38212	48	47.0	0.48852	0.33317	0.38269
24	23.0	0.51395	0.33283	0.38219	49	48.0	0.48812	0.33317	0.38270
25	24.0	0.51184	0.33287	0.38225	50	49.0	0.48761	0.33317	0.38270

We conduct yet another experiment with the dataset presented in Table 3.1. We find that ξ_1 for X_1 has the representation power 0.661265 (eigenvalue=2.645058) while ξ_2 for X_2 has the representation power 0.752979 (eigenvalue=3.764895). The $r(\xi_1,\xi_2)$ is 0.922764. Against these, results of the constrained canonical correlation (for λ=49) are: mean squared $r(\varsigma_1,x_1)$= 0.661261; mean squared $r(\varsigma_2,x_2)$=0.752966 and the constrained canonical correlation, $r(\varsigma_1,\varsigma_2)$= 0.923647. These results are once again corroborative to our earlier findings.

Table 3.1. Simulated Dataset-3 for Canonical correlation

Sl No.	X_1 or Dataset-1 X_{11}	X_{12}	X_{13}	X_{14}	X_2 or Dataset-2 X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	Sl No.	X_1 or Dataset-1 X_{11}	X_{12}	X_{13}	X_{14}	X_2 or Dataset-2 X_{21}	X_{22}	X_{23}	X_{24}	X_{25}
1	0.7	1.1	1.6	1.3	1.1	0.1	2.7	1.5	1.8	16	2.1	2.7	2.7	1.8	1.8	0.9	6.4	1.6	2.3
2	1.3	1.2	0.8	1.0	0.3	1.2	1.1	0.4	0.3	17	2.2	2.9	2.6	2.3	3.0	3.0	6.9	1.9	2.8
3	1.7	2.5	1.8	1.1	1.9	2.3	5.5	2.6	2.1	18	2.6	1.9	2.8	2.2	2.5	2.8	5.0	2.4	3.0
4	2.9	2.4	1.9	2.8	2.7	2.2	4.9	2.3	1.9	19	1.6	1.3	2.4	3.0	1.7	2.1	4.3	2.0	2.9
5	0.4	1.0	0.2	0.9	0.5	2.5	1.7	1.0	0.1	20	1.9	0.9	2.9	1.9	1.5	2.0	3.6	2.1	1.4
6	0.6	0.4	2.2	0.4	1.0	0.8	2.4	1.7	0.4	21	1.5	0.2	0.4	0.7	1.3	1.6	0.8	0.2	0.9
7	0.8	0.1	0.7	0.8	0.1	0.2	0.4	0.1	0.5	22	1.8	0.5	1.1	0.5	1.2	1.4	2.1	0.8	1.0
8	2.3	2.8	3.0	2.6	2.6	2.9	6.7	2.5	2.7	23	1.4	0.7	0.5	1.6	0.4	1.9	2.6	2.2	1.5
9	1.2	2.0	0.9	1.7	2.4	0.7	3.2	1.8	2.0	24	1.0	1.6	0.3	0.1	0.7	1.1	1.2	0.3	0.7
10	2.4	2.1	2.5	2.5	2.1	1.5	3.9	2.7	1.7	25	0.5	1.8	1.4	2.7	0.2	1.8	3.3	1.3	1.3
11	0.2	0.6	0.1	1.5	1.4	0.4	0.6	0.5	0.2	26	3.0	2.6	2.3	2.4	2.0	1.7	5.1	3.0	2.2
12	2.0	1.5	0.6	0.3	0.8	0.6	1.5	0.6	1.6	27	2.5	1.4	1.3	2.1	2.3	2.6	3.8	2.9	2.4
13	0.9	0.3	1.7	2.0	1.6	0.5	2.5	0.9	0.8	28	0.3	2.2	1.2	0.2	0.9	1.3	2.3	1.1	0.6
14	2.7	3.0	2.1	2.9	2.8	2.7	7.0	2.8	2.6	29	2.8	2.3	2.0	1.4	2.9	2.4	5.8	1.4	2.5
15	1.1	0.8	1.0	1.2	0.6	0.3	2.0	1.2	1.2	30	0.1	1.7	1.5	0.6	2.2	1.0	4.2	0.7	1.1

Fig-3. Relationship between Constrained Canonical Correlation and Sum of Squared Correlation between Canonical Variates and their Constituent Variables for Different Values of λ (Dataset in Table-3.1)

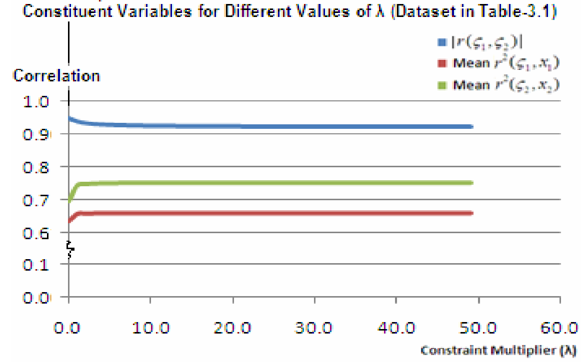


Table 3.2. Relationship between Constrained Canonical Correlation and Representation Correlation between Canonical Variates and their Constituent Variables for Different Values of λ (Dataset in Table-3.1)

Sl No.	λ	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Squared $r(\varsigma_1,x_1)$	$r(\varsigma_2,x_2)$	Sl No.	λ	Canonical $\lvert r(\varsigma_1,\varsigma_2)\rvert$	Mean Squared $r(\varsigma_1,x_1)$	$r(\varsigma_2,x_2)$
1	0.0	0.94813	0.63711	0.69854	26	25.0	0.92443	0.66125	0.75293
2	1.0	0.93901	0.65895	0.74485	27	26.0	0.92437	0.66125	0.75294
3	2.0	0.93449	0.66036	0.74955	28	27.0	0.92431	0.66125	0.75294
4	3.0	0.93199	0.66076	0.75106	29	28.0	0.92426	0.66125	0.75294
5	4.0	0.93039	0.66094	0.75175	30	29.0	0.92421	0.66125	0.75294
6	5.0	0.92927	0.66104	0.75212	31	30.0	0.92417	0.66126	0.75295
7	6.0	0.92844	0.66110	0.75235	32	31.0	0.92412	0.66126	0.75295
8	7.0	0.92780	0.66114	0.75250	33	32.0	0.92409	0.66126	0.75295
9	8.0	0.92729	0.66116	0.75260	34	33.0	0.92405	0.66126	0.75295
10	9.0	0.92687	0.66118	0.75267	35	34.0	0.92401	0.66126	0.75295
11	10.0	0.92653	0.66119	0.75272	36	35.0	0.92398	0.66126	0.75295
12	11.0	0.92624	0.66121	0.75276	37	36.0	0.92394	0.66126	0.75296
13	12.0	0.92598	0.66121	0.75279	38	37.0	0.92392	0.66126	0.75296
14	13.0	0.92577	0.66122	0.75282	39	38.0	0.92389	0.66126	0.75296
15	14.0	0.92558	0.66123	0.75284	40	39.0	0.92386	0.66126	0.75296
16	15.0	0.92541	0.66123	0.75286	41	40.0	0.92383	0.66126	0.75296
17	16.0	0.92526	0.66123	0.75287	42	41.0	0.92381	0.66126	0.75296
18	17.0	0.92513	0.66124	0.75288	43	42.0	0.92378	0.66126	0.75296
19	18.0	0.92501	0.66124	0.75289	44	43.0	0.92376	0.66126	0.75296
20	19.0	0.92490	0.66124	0.75290	45	44.0	0.92374	0.66126	0.75296
21	20.0	0.92481	0.66124	0.75291	46	45.0	0.92372	0.66126	0.75296
22	21.0	0.92472	0.66125	0.75291	47	46.0	0.92370	0.66126	0.75296
23	22.0	0.92464	0.66125	0.75292	48	47.0	0.92368	0.66126	0.75297
24	23.0	0.92455	0.66125	0.75292	49	48.0	0.92366	0.66126	0.75297
25	24.0	0.92450	0.66125	0.75293	50	49.0	0.92365	0.66126	0.75297

6. A Computer Program for RCCCA

We developed a computer program in FORTRAN that we have developed and used for solving the problems in this paper (codes available at www.webng.com/economics/rccca.txt, which may also be obtained from the author on request). Its main program (RCCCA) is assisted by 13 subroutines. The user needs setting the parameters in the main program as well as in the subroutines CORD and DORANK. Parameter setting in RPS may seldom be required. This program can be used for obtaining Ordinal Canonical Correlation [Mishra, (2009)] also. Different schemes of rank-ordering may be used [Wikipedia, (2008)].

7. Concluding Remarks

Our proposed Representation-Constrained Canonical correlation (RCCCA) Analysis has the classical canonical correlation analysis (CCCA) at its one end (λ=0) and the Classical Principal Component Analysis (CPCA) at the other (as λ tends to be very large). In between it gives us a compromise solution. By a proper choice of λ, one can avoid hijacking of the representation issue of two datasets by a lone couple of highly correlated variables across those datasets. This advantage of the RCCCA over the CCCA deserves a serious attention by the researchers using statistical tools for data analysis. Our method also addresses the problem raised by Sugiyama (2007).

8. References

[1] Eberhart, R.C. and Kennedy, J., (1995), *A New Optimizer using Particle Swarm Theory*, Proceedings Sixth Symposium on Micro Machine and Human Science*:* 39–43. IEEE Service Center, Piscataway, NJ.

[2] Fleischer, M., (2005), *Foundations of Swarm Intelligence: From Principles to Practice*, Swarming Network Enabled C4ISR, arXiv:nlin.AO/0502003 v1.

[3] Hayek, F. A., (1948), *Individualism and Economic Order*, The University of Chicago Press, Chicago.

[4] Hayek, F. A., (1952), *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, University of Chicago Press, Chicago.

[5] Hotelling, H., (1936), *Relations between Two Sets of Variates*, in: *Biometrica*, 28: 321-377.

[6] Mishra, S.K., (2006), *Global Optimization by Differential Evolution and Particle Swarm Methods: Evaluation on Some Benchmark Functions*, available at SSRN: http://ssrn.com/abstract=933827

[7] Mishra, S.K., (2009), *A Note on the Ordinal Canonical Correlation Analysis of Two Sets of Ranking Scores*. Available at *SSRN*: http://ssrn.com/abstract=1328319

[8] Sugiyama, T., (2007), *Tests for Validity on Principal Canonical Correlation Analysis*, Atlas Conference, Interdisciplinary Mathematical & Statistical Techniques (Shanghai 2007), May 20-23, 2007, University of Science and Technology of China, Hefei, Anhui, P.R. China.

[9] Urfalioglu, O., (2004), *Robust Estimation of Camera Rotation, Translation and Focal Length at High Outlier Rates*, Proceedings of the 1st Canadian Conference on Computer and Robot Vision, IEEE Computer Society Washington, DC, USA: 464- 471.

[10] Wikipedia, (2008), *Ranking*, available at Wikipedia: http://en.wikipedia.org/wiki/Rank_order