3/4/2010 36-402/608 ADA-II H. Seltman Breakout #15 Results

Summary: To perform a canonical correlation analysis in R, use cancor(X, Y) which produces a list with components "cor" (correlations of the canonical variable pairs), "xcoef" and "ycoef" (coefficients transforming data to canonical variables), and "xcenter" and "ycenter" (original X and Y column means). There is no special print() or other methods, but running cancor() does a special printout. The built in function does not compute p-values; use p.asym() (or p.perm() in package "CCP" to get the p-values.

The "University Data Set" at the UCI Machine Learning Repository has records on college and university characteristics, presumably collected in the 1980s. We will look at data on 242 institutions, considering several objective meaures of who is admitted as the explanatory variables. Our outcomes are three (subjective) quality measures. We will use CCA to find dimensions of university characterisitics that predict (correlate with) the measures of quality.

na	ames(univ)[xvar]													
#	[1] "maleFemaleRatio"	"sat.verbal"	"sat.math"	"percent.financial.aid"										
#	[5] "percent.admittanc	e" "percent.enrolled"												
names(univ)[yvar]														
#	[1] "academics.scale"	"social.scale"	"quality.of.life.scale"											
r	round(cor(univ[vuar])) 2)													
#		academics.scale social	L.scale qua	Lity.of.life.scale										
#	maleFemaleRatio	-0.04	-0.01	0.11										
#	sat.verbal	-0.03	-0.04	0.16										
#	sat.math	0.02	0.03	0.08										
#	percent.financial.aid	-0.12	-0.13	-0.14										
#	percent.admittance	0.11	-0.05	-0.11										
#	percent.enrolled	0.07	0.07	0.06										

			-								
0	0	0	0	0	0				0	- 6	
	 o@@@@@	666666 60	6000 0	09 9799	6000000000000000000000000000000000000	b 🍘 🍯	9 68 6	D 📕 👁	•	- 5	4 5 6
0 0 0 0		2000000			°	0				ality.of	.life_sca
										1 2 3	3 -
							9 0 A	9 9 9 9 0			2 -
000000			0,000 000		6885686	jų č	998	- 5 I	Ϋ́	880	
6		🛞 o 🦲		A a b	900 / D	•	0 0	4 3	4 5	ə 👩	
88°00			0000	888	0000	00	20	social	scale	໌ 🔬 ິ ຨ	\sim
			ക്കുറുള്ളം. ത		ഞ്ഞെറ്റെപ്പ		8 8 8 8 10 10	1 2 2	2 _		
								1 2 3	~	, U T	
<u>an</u> 0	8 6 6 8 20					- 5		0 0	0.6		8
0	BOR R Q		1 00 00	888 6 8 0	000,000	- 4 3	4 5	0 🔊 90	0.6	o 🖗 🌰	. 🙆
						ademi	rs sral			0.0	Ă
						1 2 3				200	
	HOW HE AND A			00 00							
ക്രഹം					- ¹⁵ 10 15	ີ່	í.	° 8 -	Ø	ୁହି	O A
8000						യം ത്ര	800				କୁ
	Sem and a series of the series						Ö Ö	₿₿00	o 📆	3 ° 🚆	8 B
A 8000) 5 5 -		6 6			2	<u>`````````````````````````````````````</u>
			0000				<u> </u>			<u> </u>	
6 0 8. °Q	ളം രം ക	88.00.00	മ എം വാ	15 10 15 3	ക്രാംസ്	2 ~ 6	0	୶ୖୖୖ୶ୖ		Ĭ 🦓	8
				10 10 10 2			0		0	<u> </u>	<u>م </u>
	2 2 ~ ⁄~			rcent.admittar		° 👷 🖉	<u> </u>	8 🚨	_ 😤	రి 🙀	
		0 00000		5 10 5 -				8	0 2	P	0
<u>6000000000000000000000000000000000000</u>	and a second second		a source			9 🕒 🤅	,	9 8 6		368	
80000			- 20	<mark>ക</mark> ം റ്റ്റ്	പ്പ് ം പ്	S_{2}	່ຂັ	്ക	8	್ಗೆ ೧ &	R
80 00 O	ഁ൙ൣ൙	്ക്ക്റ്റ്	- 15 ¹⁰ 15 ²⁰	\$ 8,800	66000 016		6° c	ŏŏo	ģ	த் இ	ğ
	စို့စွဲစို ထင်ဝ	600 00 0 0 0 0	centinangiaL	000 0 0000		5 92 9		စ တို့ ၀	0	ະ 🖁	🦉 🧕 🖉
	Contraction of the second		05105-				jõ 🦉		- e 🍂		8
\$%\$%	6 66	8968968		6 66	0000		° 🏹			§ 🖲 🦉	~
®oo ₀°	8 o 🔗		6 0000	00000	°°° ~	88	0	° 9	8	8	0
lo l	് െ പ്	$\begin{bmatrix} 50 & 30 & 40 & 50 & 60 \\ 40 & & & & & & & \\ \end{bmatrix}$				ି ଭୁଇ ପ୍ରି	8 0 0	ຄິວຈ		്രജ്	2
	a la coo	– sat.mash –	Be o o B				ξσο	ັ 🌢 🖏		စုစု 🚳	
		0 10 20 30 20 -	အြင္ရ၀ ရွ င္လဲ တစ ္စင				8 @ @	a a à	0	2 8 🚳	A 1
	2		and the second			x 2 3	8	ð 👗 à		18 8	V
0 8 08		00 000	\$ 0 0 °	.00 A	<u>a</u> 8	6	20	00	ø	6	0
	50 40 30 40 50 6		କ୍ଷିତ୍ର ଦିବ୍ଦ ଦ			ີ່ຄິດເອ	3 ັ	8		, 🍃	6
86 28 0	— s39tver1864 —		6 6 °		ങ്കെ 8 ഹം പ	\$ 8 6		000			· 🔗
	20 -			യ്ക്കായില് അ			o e	o xo 6	o 🎽	ှိဝ 🍓	
	0 10 20 30 10 -								2		
) O	
20 30		ര്റ്റ്റ്							a de la comercia de l	, &	ര
	a la			8000	ိုလ်လို	ູຈີ່ຊີ່ຊີ	° 0	8.		ည်ရှိ	8
alereinalekat	20088ggg	8 88 88 o	Back Book	36	8 ° ° ° ° °		38.	6 & &	_ %	ខ្មុំ 凝	8
0 10 10 -			00000					8 🔹 👳	0,	jõ h	
	0000 00 0000	COLUMNOUS COLUMNOUS	0 9660000				900			, 🚳 🛑	· · · •

University Data

Question 1: What do you observe on the original scales?

```
cc=cancor(univ[,xvar],univ[,yvar])
print(round(cc$cor,3))
# 0.271 0.257 0.111
print(round(cc$xcoef,4))
                           [,1]
                                   [,2]
                                           [,3]
                                                   [,4]
                                                           [,5]
                                                                   [,6]
#
# maleFemaleRatio
                         0.0022 0.0031 -0.0001 -0.0021 -0.0098 0.0005
# sat.verbal
                         0.0010 0.0043 -0.0043 0.0022
                                                         0.0022 -0.0002
# sat.math
                         0.0007 -0.0017
                                        0.0026 -0.0057
                                                         0.0011
                                                                 0.0014
# percent.financial.aid -0.0128 0.0070 0.0034 0.0002
                                                         0.0005
                                                                 0.0069
# percent.admittance
                        -0.0034 -0.0107 -0.0138 -0.0025 -0.0036
                                                                 0.0003
# percent.enrolled
                         0.0096 -0.0062 0.0063 0.0104 -0.0009
                                                                 0.0134
```

print(round(cc\$ycoef,3)) [,1] [,2] [,3] # #academics.scale 0.025 -0.059 -0.047 #social.scale 0.021 -0.020 0.041 #quality.of.life.scale 0.046 0.035 -0.015 require(CCP) p.asym(cc\$cor, nrow(univ), length(xvar), length(yvar)) #Wilks' Lambda, using F-approximation (Rao's F): # stat approx df1 df2 p.value #1 to 3: 0.8551364 1.2977101 18 410.6072 0.1846153 #2 to 3: 0.9226508 1.1993408 10 292.0000 0.2908825 #3 to 3: 0.9877572 0.4554981 4 147.0000 0.7682613

Question 2: What tells you that we haven't found any interesting new scales? If you pretend that the first p-value is 0.00185, what conclusions would you reach?