Tema 4

Análisis de Correlación Canónica

4.1. Introducción

El análisis de correlación canónica es una técnica multivariante que actualmente se usa para analizar relaciones multidimensionales entre múltiples variables independientes métricas y múltiples variables dependientes métricas.

Inicialmente el análisis de correlación canónica se utilizó para investigar si existía alguna relación entre dos grupos de variables. Hotteling [6], en 1936, fue el primero en desarrollarlo. Su interés se centró en describir los resultados de las pruebas para determinar la velocidad de lectura X_1, la capacidad de compresión X_2, la velocidad de cálculo aritmético Y_1 y la precisión aritmética Y_2, obtenidos en una muestra de 140 estudiantes de secundaria. Su objetivo específico fue analizar si la habilidad de lectura (medida por X_1 y X_2) estaba relacionada con la habilidad aritmética (medida por Y_1 e Y_2).

El anterior objetivo puede estudiarse vía el análisis de correlación canónica. Para ello se necesitan construir combinaciones lineales de X_1 y X_2,

$$U = a_1X_1 + a_2X_2,$$

y combinaciones lineales de Y_1 e Y_2,

$$V = b_1Y_1 + b_2Y_2,$$

donde los coeficientes a_1, a_2, b_1 y b_2 se escogen de manera tal que la correlación entre U y V sea lo más grande posible. Esta metodología es similar en cierto sentido al análisis de componentes principales, con la diferencia de que en este análisis se busca maximizar la correlación en lugar de la variabilidad.

Hotteling encontró que

$$U = -2.78X_1 + 2.27X_2 \quad \text{y} \quad V = -2.44Y_1 + Y_2,$$

con una correlación entre ellas de 0.62, demostrando que estudiantes con diferencias entre la velocidad de lectura y su comprensión, tienden a tener diferencias entre su velocidad aritmética y su precisión. Este aspecto determinado por U y V es el que más correlación da entre la habilidad de lectura y la aritmética.

El análisis de correlación canónica se aplica hoy en día a situaciones donde es apropiada la técnica de la regresión pero existe más de una variable dependiente. Otra aplicación del análisis de correlación canónica, volviendo a su origen, es como un método para determinar la asociación entre dos grupos de variables. Aplicaciones a la psicología se pueden encontrar en Cooley y Lohnes [1], Cuadras y Sánchez [3]. En ecología se ha aplicado como un modelo para estudiar la relación entre presencia de especies y variables ambientales (ver Gittings [4]). La distribución de las correlaciones canónicas es bastante complicada y solamente se conocen resultados asintóticos (Ver Muirhead [8]).

Esta técnica estadística multivariante esta íntimamente relacionada con otras como el análisis canónico discriminante, que se estudiará más adelante, y tiene ciertas propiedades análogas al análisis de componentes principales y al análisis factorial pero, en lugar de tratar de estudiar las dependencias internas entre las variables de un mismo grupo, lo que se estudia es la relación o dependencia entre dos grupos de variables, unas dependientes y otras independientes. En general, como se ha indicado previamente, se aplica a situaciones donde es apropiada la técnica de regresión múltiple pero para más de una variable dependiente cuando dichas variables muestran correlación entre sí. Si las variables dependientes fueran incorreladas entre sí, se podría optar por aplicar una regresión lineal múltiple a cada una de ellas pero, en caso contrario, como acaba de indicarse, es la técnica de correlación canónica la respuesta al estudio.

Uno puede plantearse alternativamente utilizar el análisis factorial en cada grupo de variables ya que, como ya se ha visto en un tema previo, esta técnica se caracteriza por construir variables que son combinación lineal de las originales, de modo que se maximiza el poder explicativo de las causas comunes que ocasionan la variabilidad del conjunto de variables. Además, este análisis garantiza que las variables generadas por este procedimiento están incorreladas entre sí, lo que asegura la inexistencia de multicolinealidad entre estas variables ficticias, pudiendo ser utilizadas como independientes en un modelo de regresión múltiple y, a su vez, al aplicarlo al conjunto de variables dependientes, se obtendría un conjunto de variables dependientes ficticias incorreladas entre sí que permitiría aplicar, como se ha indicado previamente, una regresión lineal múltiple para cada una de ellas. El inconveniente de este procedimiento es que no se garantiza que estas variables ficticias independientes conserven el poder explicativo sobre las variables ficticias dependientes.

Por lo tanto, sería deseable la aplicación de una técnica estadística multivariante que generara un conjunto reducido de variables dependientes, y otro de variables independientes, de manera que ambos grupos estuviesen altamente correlacionados, mientras que las causas comunes entre las variables que forman parte de cada grupo fuesen nulas. El análisis de correlación canónica proporciona variables que garantizan este resultado y es recomendable su uso cuando tanto el número de variables dependientes como independientes sea elevado y existan elevadas correlaciones de las variables de cada grupo entre sí. Mediante este método de reducción de variables se puede eliminar el ineludible problema de multicolinealidad que ocasionaría la consideración de todas las variables en un modelo de regresión.

En general la técnica consiste en encontrar una combinación lineal de las variables $X = (X_1, X_2, \ldots, X_p)$,

$$U_1 = Xa_1 = a_{11}X_1 + a_{12}X_2 + \ldots + a_{1p}X_p,$$

y otra combinación lineal de las variables $Y = (Y_1, Y_2, \ldots, Y_q)$,

$$V_1 = Yb_1 = b_{11}Y_1 + b_{12}Y_2 + \ldots + b_{1q}Y_q,$$

siendo $a_1 = (a_{11}, \ldots, a_{1p})'$ y $b = (b_{11}, \ldots, b_{1q})'$, de tal manera que la correlación entre U_1 y V_1 sea máxima. Después encontrar otras dos combinaciones lineales para cada grupo de variables que tenga correlación máxima y así sucesivamente se encuentran un conjunto de combinaciones lineales para cada grupo de variables que tienen correlación máxima. A estas combinaciones lineales se denominan variables canónicas y las correlaciones entre los correspondientes pares de variables canónicas se denominan correlaciones canónicas.

Hasta hace pocos años el análisis de correlación canónica era una técnica estadística relativamente desconocida. La disponibilidad de programas de computación ha facilitado el aumento de su utilización en problemas de investigación. Aún así el análisis de correlación canónica es, de las técnicas estadísticas multivariantes, uno de los menos utilizados debido, en parte, a la dificultad que se puede encontrar a la hora de interpretar los resultados.

Hoy en día el análisis de correlación canónica se ha generalizado al caso en que hay más de dos grupos de variables y también existe la versión no lineal de esta técnica, pero en este tema nos centraremos en el estudio de la técnica clásica.

4.2. Cálculo de variables canónicas

Siguiendo a Cuadras [2], se consideran dos conjuntos de variables: una serie de variables independientes $(X_1, X_2, \ldots, X_p)$ y otra serie de variables dependientes $(Y_1, Y_2, \ldots, Y_q)$ sobre un grupo de objetos o individuos y se trata de calcular, a partir de ellas, dos nuevos conjunto de variables $(U_1, U_2, \ldots, U_m)$ y $(V_1, V_2, \ldots, V_m)$ que verifiquen que:

- $Corr[U_1, V_1] \geq Corr[U_2, V_2] \geq \ldots \geq Corr[U_m, V_m]$ máxima.
- $Corr[U_i, U_j] = 0 \ \forall i \neq j$ (Las variables canónicas $U_1, U_2, \ldots, U_m$ están incorreladas).
- $Corr[V_i, V_j] = 0 \ \forall i \neq j$ (Las variables canónicas $V_1, V_2, \ldots, V_m$ están incorreladas).
- $Corr[U_i, V_j] = 0 \ \forall i \neq j$.

4.2.1. Obtención de las primeras variables canónicas

Las primeras componentes (U_1, V_1) se calculan eligiendo vectores $a_1 = (a_{11}, \ldots, a_{1p})'$ y $b_1 = (b_{11}, \ldots, b_{1q})'$ de modo que $Corr[U_1, V_1]$ sea la mayor posible. Denotemos por S_{XX} y S_{YY} las matrices de covarianzas muestrales de los vectores X e Y y sea $S_{XY} = S'_{YX}$ la matriz de covarianzas muestrales entre el vector X y el Y. Puesto que la correlación depende a su vez de las varianzas de U_1 y V_1, ambas no acotadas, se va a imponer la restricción de que dichas varianzas sean 1. Por tanto, queremos elegir a_1 y b_1 de modo que se maximice

$$Corr[U_1, V_1] = \mathrm{Cov}[U_1, V_1] = \mathrm{Cov}[Xa_1, Yb_1] = a_1' \,\mathrm{Cov}[X, Y]b_1 = a_1' S_{XY} b_1$$

sujeta a las restricciones de que

$$\text{Var}[U_1] = \text{Var}[Xa_1] = a_1' \,\text{Var}[X]a_1 = a_1' S_{XX} a_1 = 1 \text{ y } b_1' S_{YY} b_1 = 1.$$

Aplicando el método de los multiplicadores de Lagrange, tenemos

$$L(a_1, b_1) = a_1' S_{XY} b_1 - \frac{\lambda}{2}(a_1' S_{XX} a_1 - 1) - \frac{\delta}{2}(b_1' S_{YY} b_1 - 1)$$

y buscamos el máximo derivando e igualando a 0

$$\frac{\partial L}{\partial a_1} = S_{XY} b_1 - \lambda S_{XX} a_1 = 0,$$

$$\frac{\partial L}{\partial b_1} = S_{YX} a_1 - \delta S_{YY} b_1 = 0.$$

Multiplicando la primera ecuación por a_1' y la segunda por b_1' tenemos:

$$a_1' S_{XY} b_1 - \lambda a_1' S_{XX} a_1 = a_1' S_{XY} b_1 - \lambda = 0,$$

$$b_1' S_{YX} a_1 - \delta b_1' S_{YY} b_1 = b_1' S_{YX} a_1 - \delta = 0.$$

Ahora, teniendo en cuenta que $a_1' S_{XY} b_1 = b_1' S_{YX} a_1$, se concluye que

$$\lambda = \delta = a_1' S_{XY} b_1 = Corr[U_1, V_1].$$

Por otro lado, de la segunda ecuación del sistema de ecuaciones normales se tiene (siempre que la matriz de covarianzas de Y sea no singular o invertible)

$$b_1 = \frac{1}{\lambda} S_{YY}^{-1} S_{YX} a_1$$

y sustituyendo en la primera ecuación se llega a $S_{XY} S_{YY}^{-1} S_{YX} a_1 (\lambda)^{-1} - \lambda S_{XX} a_1 = 0$. Multiplicando ahora esta ecuación por S_{XX}^{-1} (de nuevo asumiendo que existe dicha inversa) y λ se obtiene

$$S_{XX}^{-1} S_{XY} S_{YY}^{-1} S_{YX} a_1 - \lambda^2 I a_1 = (S_{XX}^{-1} S_{XY} S_{YY}^{-1} S_{YX} - \lambda^2 I) a_1 = 0,$$

es decir, $Corr[U_1, V_1] = \lambda_1$ siendo λ_1^2 un valor própio de $S_{XX}^{-1} S_{XY} S_{YY}^{-1} S_{YX}$.

Por tanto, para maximizar la correlación entre las primeras variables canónicas se debe seleccionar al mayor autovalor de la matriz indicada. Además, el primer vector canónico a_1 será el autovector asociado a dicho mayor autovalor verificando la restricción $a_1' S_{XX} a_1 = 1$. Para obtener el vector canónico b_1 se puede usar la ecuación obtenida $b_1 = S_{YY}^{-1} S_{YX} a_1 / \lambda$ imponiéndole también la restricción $b_1' S_{YY} b_1 = 1$.

Todo el desarrollo realizado para determinar a_1 y, apartir de él, determinar b_1 puede repetirse cambiando los papeles llegándose a que b_1 debe ser el vector propio asociado al mayor valor propio de la matriz $S_{YY}^{-1} S_{YX} S_{XX}^{-1} S_{XY}$ con la restricción $b_1' S_{YY} b_1 = 1$ y a_1 se puede obtener de la ecuación $a_1 = S_{XX}^{-1} S_{XY} b_1 / \lambda$ con la restricción $a_1' S_{XX} a_1 = 1$.

Ya que tanto los a_i como los b_i, es decir los vectores canónicos, deben cumplir las dos resoluciones anteriores, se pueden obtener $m = \text{Min}(p, q)$ pares de variables canónicas a partir de X e Y siempre que se obtengan m valores propios distintos de la matriz $S_{XX}^{-1} S_{XY} S_{YY}^{-1} S_{YX}$ o, equivalentemente, $S_{YY}^{-1} S_{YX} S_{XX}^{-1} S_{XY}$. Cada valor propio y su correspondiente vector propio determinarán una correlación canónica y unas variables canónicas. Siempre se debe empezar por el mayor valor propio, que será el que origine una correlación canónica mayor, e ir decreciendo si se quieren buscar más correlaciones canónicas.

4.2.2. Correlación canónica y descomposición singular

Podemos formular una expresión conjunta para los vectores canónicos utilizando la descomposición singular de una matriz.

Supongamos $p > q$; consideremos la matriz $p \times q$

$$Q = S_{XX}^{-1/2} S_{XY} S_{YY}^{-1/2}$$

y hallemos $Q = U \Lambda V'$, la descomposición singular de Q, donde U es una matriz $p \times q$ con columnas ortonormales, V es una matriz $q \times q$ ortogonal, y Λ es una matriz diagonal con los valores singulares de Q. Es decir,

$$U'U = I_q,\ V'V = VV' = I_q,\ \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_q).$$

Entonces los vectores canónicos y correlaciones canónicas son:

$$a_i = S_{XX}^{-1/2} u_i, \quad b_i = S_{YY}^{-1/2} v_i, \quad Corr[U_i, V_i] = \lambda_i$$

Se puede probar que las correlaciones canónicas son invariantes por transformaciones lineales y, en consecuencia, pueden calcularse a partir de las matrices de correlaciones en lugar de a partir de las matrices de covarianzas muestrales como se ha realizado aquí.

Ejemplo: se toma una muestra de diez individuos (caprinos) en los que se miden dos variables productivas de leche, como pueden ser producción máxima diaria (Y_1) y porcentaje de nitrógeno total (Y_2), y dos variables de conformación, como pueden ser longitud total del cuerpo (X_1) y anchura de las caderas (X_2). Los datos son

Individuo	Y_1	Y_2	X_1	X_2
1	122	40	332	116
2	120	42	320	107
3	126	44	339	119
4	125	39	336	114
5	120	38	321	106
6	127	45	336	119
7	128	49	347	128
8	130	39	349	129
9	123	41	338	111
10	124	42	333	112

En este ejemplo vamos a trabajar a partir de la matriz de correlaciones:

$$R = \begin{pmatrix} 1 & 0.908742 & 0.92525 & 0.38379 \\ 0.90874 & 1 & 0.92782 & 0.46868 \\ 0.92525 & 0.92782 & 1 & 0.4121 \\ 0.38379 & 0.46868 & 0.41212 & 1 \end{pmatrix}.$$

Entonces:

$$R_{XX} = \begin{pmatrix} 1 & 0.90874 \\ 0.90874 & 1 \end{pmatrix} \quad R_{YY} = \begin{pmatrix} 1 & 0.41212 \\ 0.41212 & 1 \end{pmatrix}$$

$$R_{XY} = \begin{pmatrix} 0.92525 & 0.38379 \\ 0.92782 & 0.46868 \end{pmatrix} \quad R_{YX} = \begin{pmatrix} 0.92525 & 0.92782 \\ 0.38379 & 0.46868 \end{pmatrix}.$$

Calculamos ahora la matriz a la que debemos determinar sus valores y vectores propios para obtener las correlaciones canónicas:

$$R_{XX}^{-1} R_{XY} R_{YY}^{-1} R_{YX} = \begin{pmatrix} 0.877559 & 0.385797 \\ 0.0533429 & 0.0708581 \end{pmatrix}. \tag{4.1}$$

A partir de esta matriz, obtenemos la ecuación característica cuyas raíces nos darán los valores propios:

$$|R_{XX}^{-1} R_{XY} R_{YY}^{-1} R_{YX} - \lambda^2 I_2| = \lambda^4 - 0.9484171\lambda^2 + 0.04160260 = 0$$

En este caso, los valores propios son $\lambda_1^2 = 0.90231$ y $\lambda_2^2 = 0.04611$, y de ahí podemos obtener las correlaciones canónicas tomando sus raíces cuadras: $\lambda_1 = 0.9499$ y $\lambda_2 = 0.2147$.

El primer vector propio, asociado al valor propio mayor, nos daría el primer vector canónico a_1 al normalizarlo, es decir, al imponerle la restricción $a_1' S_{XX} a_1 = 1$. Para determinarlo resolvemos el sistema:

$$(R_{XX}^{-1} R_{XY} R_{YY}^{-1} R_{YX} - \lambda_1^2 I_2) v_1 = 0$$

obteniendo

$$v_1 = (0.642463, 0.766316)' \Rightarrow v_1' S_{XX} v_1 = 132.827 \Rightarrow a_1 = \frac{1}{\sqrt{132.827}} v_1 = (0.04947648, 0.07077162)'.$$

Para obtener el vector canónico b_1 tenemos dos opciones. La primera sería repetir el procedimiento anterior a partir de la matriz

$$R_{YY}^{-1} R_{YX} R_{XX}^{-1} R_{XY} = \begin{pmatrix} 0.434787 & 0.391962 \\ 0.463609 & 0.513630 \end{pmatrix}. \tag{4.2}$$

Sus valores propios siguen siendo los mismos de antes pero el primer vector propio, asociado al valor propio mayor 0.90231, sí que cambia. Ahora tenemos

$$v_1 = (0.997948, 0.0640247)' \Rightarrow v_1' S_{YY} v_1 = 11.6972 \Rightarrow b_1 = \frac{1}{\sqrt{11.6972}} v_1 = (0.2910728, 0.0186371)'.$$

La otra opción es usar la fórmula que relaciona b_1 con a_1: $b_1 = S_{YY}^{-1} S_{YX} a_1 / \lambda$.

Por lo tanto, las primeras variables canónicas son:

$$U_1 = 0.049476482X_1 + 0.0707716X_2 \quad V_1 = 0.2910728Y_1 + 0.0186371Y_2,$$

con correlación canónica $\lambda_1 = 0.9499$.

Un método para interpretar el valor relativo de cada variable en la combinación lineal canónica, es viendo el valor de los coeficientes de los vectores canónicos. Así, para las Y, la primera variable canónica viene determinada fundamentalmente por la variable producción, lo que quiere decir que un individuo que produzca relativamente mucho tendrán un alto valor de la primera variable canónica V_1. Mientras, para las X, en el valor de la primera variable canónica U_1, aunque tiene una influencia ligeramente superior la anchura que la longitud, se puede considerar que la influencia es la misma. El problema de esta interpretación es que asume que las unidades de medida o su escala es la misma para todas las variables bajo estudio, lo cual no tiene por qué ser cierto. Por ello, para interpretar el valor relativo de cada variable en la combinación lineal canónica, es mejor usar el valor de la correlación de cada variable original con su variable canónica (o con la variable canónica del otro grupo de variables), ya que el coeficiente de correlación es adimensional, es decir, no depende de la unidad de medida que tenga la variable.

R	U_1	V_1
X_1	0.9726324	0.9239037
X_2	0.9808436	0.9317036
Y_1	0.9483636	0.9983823
Y_2	0.4400446	0.4632535

Como se ve, la primera variable canónica de las Y (V_1) está altamente correlacionada con la Y_1 y medianamente correlacionada con la Y_2 por lo que se puede determinar que la variable canónica V_1 viene determinada fundamentalmente por la variable producción, lo que quiere decir que un individuo que produzca relativamente más, tendrán un alto valor de la variable canónica V_1. Mientras que en el valor de la primera variable canónica de las X (U_1) tiene una correlación ligeramente superior con anchura que la longitud, pero en ambas es elevada. De todo esto se deduce que los individuos muy anchos y largos tienen una elevada producción pero no indica nada de la proporción. Lógicamente, en un ejemplo real con más variables el análisis de estos coeficientes puede ser gran utilidad para conducir a múltiples y variadas conclusiones.

Se pueden realizar interpretaciones adicionales de la relación entre las X y las Y obteniendo otro conjunto de variables canónicas y su correspondiente correlación canónica. Ya que $m = Min(p, q) = Min(2, 2) = 2$, se puede hallar la segunda variable canónica U_2 (combinación lineal de las X) y la correspondiente variable canónica V_2 (combinación lineal de las Y). Como hemos visto previamente, los coeficientes de estas combinaciones lineales se eligen teniendo en cuenta las siguientes condiciones: 1. U_2 esta incorrelacionada con U_1 y V_1. 2. V_2 esta incorrelacionada con U_1 y V_1. 3. Una vez cumplidas las condiciones anteriores, U_2 y V_2 tienen la máxima correlación posible. La correlación entre U_2 y V_2 se denomina segunda correlación canónica y necesariamente es menor o igual que la primera correlación canónica.

Para determinarla basta con coger el segundo valor propio de la matriz obtenida previamente (4.1), el cual ya vimos que era $\lambda_2^2 = 0.04611$, y su raíz cuadrada nos da la segunda correlación

canónica $\lambda_2 = 0.2147$. Análogamente, para determinar el segundo vector canónico, se debe determinar el segundo vector propio y normalizarlo:

$$v_2 = (-0.710073, 0.704128)' \Rightarrow v_2' S_{XX} v_2 = 14.8594 \Rightarrow a_2 = \frac{1}{\sqrt{14.8594}} v_2 = (-0.2491248, 0.2962556)'.$$

Análogamente, usando la matriz (4.2), tenemos el mismo segundo valor propio pero su segundo vector propio normalizado nos permitirá determinar b_2:

$$v_2 = (-0.420901, 0.907107)' \Rightarrow v_2' S_{YY} v_2 = 8.03428 \Rightarrow b_2 = \frac{1}{\sqrt{8.03428}} v_2 = (-0.1521514, 0.3272590)'.$$

De esta forma llegamos a que las segundas variables canónicas son:

$$U_2 = -0.2491248 X_1 + 0.2962556 X_2 \quad V_2 = -0.1521514 Y_1 + 0.3272590 Y_2,$$

con correlación canónica $r_1 = 0.2147$, y las correlaciones con las variables originales son:

R	U_2	V_2
X_1	-0.23234939	-0.04989117
X_2	0.1947967	0.04182769
Y_1	-0.01220868	-0.05685696
Y_2	0.1902947	0.88622582

En este caso se observa que en la segunda variable canónica de las Y (V_2) tiene una influencia negativa la variable Y_1, baja en valor absoluto comparada con la influencia positiva de la variables Y_2, esto significa que en la segunda variable canónica de las Y tiene una gran influencia positiva la variable porcentaje de nitrógeno y una leve influencia negativa la variable producción. Mientras que en el valor de la segunda variable canónica de las X (U_2) tienen prácticamente la misma influencia las dos variables pero en sentido contrario, esto es, la longitud tiene una influencia negativa en el valor de su segunda variable canónica y la anchura tiene una influencia positiva, pero en ambas variables esta influencia es más bien baja.

4.3. Contrastes de significación

Hemos encontrado las variables y correlaciones canónicas a partir de las matrices de covarianzas y correlaciones muestrales, es decir, a partir de muestras de tamaño n. Naturalmente, todo lo que hemos dicho vale si sustituimos S_{XX}, S_{XY} y S_{YY} por las versiones poblacionales Σ_{XX}, Σ_{XY} y Σ_{YY}.

Siguiendo a Rencher [9], estudiemos ahora una serie de contrastes de hipótesis que nos permitirán comprobar la significación de las correlaciones canónicas obtenidas, es decir que existe relación lineal entre ellas. Para ello será necesario asumir la normalidad multivariante de X e Y.

4.3.1. Contraste de hipótesis de independencia

Afirmar que el vector X es independiente del vector Y consiste en plantear

$$\begin{cases} H_0 : \Sigma_{XY} = 0 \\ H_1 : \Sigma_{XY} \neq 0 \end{cases}$$

Si $\Sigma_{XY} = 0$ entonces $\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} = 0$ y la hipótesis de independencia es equivalente a la hipótesis de que todas las correlaciones canónicas poblacionales son cero:

$$H_0 : \rho_1 = \ldots = \rho_m = 0$$

siendo

$$\rho_1 \geq \rho_2 \geq \ldots \geq \rho_m$$

las $m = \text{Min}(p, q)$ correlaciones canónicas obtenidas a partir de las soluciones de:

$$|\Sigma_{XX}^{-1}\Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} - \lambda^2| = 0.$$

Razón de verosimilitud Si la hipótesis es cierta, entonces el test de razón de verosimilitud nos proporciona el estadístico Λ de Wilks

$$\Lambda_1 = \frac{|S|}{|S_{XX}||S_{YY}|} = \frac{|R|}{|R_{XX}||R_{YY}|},$$

que sigue la distribución $\Lambda(p; n-1-q; q)$.

Es fácil probar que Λ_1 es función de las correlaciones canónicas

$$\Lambda_1 = |I - S_{YY}^{-1}S_{YX}S_{XX}^{-1}S_{XY}| = \prod_{i=1}^{m}(1-\lambda_i^2).$$

Los valores críticos para $\alpha = 0.05$ de la Λ de Wilks están tabulados pero se suelen usar otros estadísticos que se obtienen a partir de él, y que tienen distribución asintótica χ^2 o F de Snedecor, para la resolución de este contraste.

La primera opción es usar el estadístico

$$\chi_1^2 = -[n - 0.5(p+q+3)]\ln\Lambda_1 = -[n-1-0.5(p+q+1)]\sum_{i=1}^{m}\ln(1-\lambda_i^2) \overset{n\to\infty}{\underset{H_0}{\rightsquigarrow}} \chi_{qp}^2.$$

Se rechaza la hipótesis nula si $\chi_1^2 > \chi_{qp,\alpha}^2$ o si $P[\chi_{qp}^2 > \chi_1^2] < \alpha$.

La otra opción es usar el estadístico cuya distribución asintótica es:

$$F_1 = \frac{wt - \frac{1}{2}pq + 1}{pq}\frac{1-\Lambda_1^{1/t}}{\Lambda_1^{1/t}} \overset{n\to\infty}{\underset{H_0}{\rightsquigarrow}} F_{pq,wt-\frac{1}{2}pq+1}$$

donde $w = n - \frac{1}{2}(p+q+3)$ y $t = \sqrt{(p^2q^2-4)/(p^2+q^2-5)}$. Se rechaza la hipótesis nula si $F_1 > F_{pq,wt-\frac{1}{2}pq+1,\alpha}$ o si $P[F_{pq,wt-\frac{1}{2}pq+1} > F_1] < \alpha$.

Para este estadístico se tiene la ventaja de que si $m = 1$ o 2, entonces su distribución es exacta.

Existen otros estadísticos que también pueden usarse para resolver este contraste:

- El estadístico Traza de Pillai-Bartlett: $V_1 = \sum_{i=1}^{m} \lambda_i^2$.
- El estadístico de la Traza de Hotelling-Lawley: $U_1 = \sum_{i=1}^{m} \frac{\lambda_i^2}{1-\lambda_i^2}$.

En el libro de Rencher [9] están tabulados los valores críticos para todos estos tests, pero también se pueden encontrar aproximaciones de estos estadísticos que tienen distribuciones asintóticas basadas en la distribución F.

- Aproximación del estadístico Traza de Pillai-Bartlett:

$$F_{V_1} = \frac{(n-p-q+m-1)V_1}{(|p-q|+m)(m-V_1)} \overset{n\to\infty}{\underset{H_0}{\rightsquigarrow}} F_{m(|p-q|+m),m(n-p-q+m-1)}.$$

- Aproximación del estadístico Traza de Hotelling-Lawley:

$$F_{U_1} = \frac{[m(n-q-p-2)+2]}{m^2(|q-p|+m)} U_1 \overset{n\to\infty}{\underset{H_0}{\rightsquigarrow}} F_{m(|p-q|+m),m(n-p-q-2)+2}.$$

En ambos casos, como ocurría con el test de Wilks, se rechaza la H_0 cuando el valor del estadístico supere al valor de la distribución correspondientes que deje una probabilidad α a la derecha o si el p-valor es menor que α.

Si se rechaza la hipótesis nula del contraste se puede deducir que al menos una correlación es significativa y, dado el orden que hay entre ellas, es evidente que al menos la primera lo es, lo cual nos puede llevar a plantearnos si el resto lo son o no.

Ejemplo: Contrastemos la independencia entre los vectores X e Y de nuestro ejemplo, es decir, contrastemos si las variables relacionadas con la producción de leche de nuestro estudio y las variables relacionadas con la conformación de nuestros individuos son independientes. Para ello usaremos el contraste de independencia que acabamos de estudiar:

$$\begin{cases} H_0: \ X \text{ e } Y \text{ son independientes} \Leftrightarrow \Sigma_{XY} = 0 \Leftrightarrow \rho_1 = \rho_2 = 0 \\ H_1: \ X \text{ e } Y \text{ no son independientes} \Leftrightarrow \Sigma_{XY} \neq 0 \Leftrightarrow \rho_1 = 0 \text{ ó } \rho_2 = 0 \end{cases}$$

En nuestro caso, la distribución del estadístico F, una de las posibles aproximaciones de Λ de Wilks, es exacta ya que $m = \text{mín}(p,q) = \text{mín}(2,2) = 2$, por lo tanto dicho estadístico sería la mejor opción, sobre todo teniendo en cuenta que sólo disponemos de 10 datos muestrales.

$$F_1 = \frac{wt - \frac{1}{2}pq + 1}{pq} \frac{1-\Lambda_1^{1/t}}{\Lambda_1^{1/t}}$$

con

$$w = n - \frac{1}{2}(p+q+3) = 10 - \frac{7}{2} = \frac{13}{2},$$

$$t = \sqrt{\frac{p^2q^2-4}{p^2+q^2-5}} = \sqrt{\frac{16-4}{4+4-5}} = \sqrt{\frac{12}{3}} = 2,$$

$$\Lambda_1 = \prod_{i=1}^{m}(1-\lambda_i^2) = (1-\lambda_1^2)(1-\lambda_2^2) = (1-0.90231)(1-0.04611) = 0.09319.$$

Por lo tanto, el estadístico experimental queda

$$F_1 = \frac{12}{4}\frac{1-0.09319^{1/2}}{0.09319^{1/2}} = 6.82766 > F_{pq,wt-\frac{1}{2}pq+1} = F_{4,12;0.05} = 3.26$$

y se rechaza la hipótesis nula, es decir, los vectores X e Y no son independientes en base a los datos de que disponemos y a un nivel de significación de 0.05.

Si optamos por usar cualquiera de los otros estadísticos estudiados, estos serían los resultados:

- Aproximación de Λ de Wilks basada en la distribución χ^2:

$$\begin{aligned}\chi_1^2 =& -[n-0.5(p+q+3)]\ln\Lambda_1 = \\ =& -[10-0.5\cdot 7][\ln(1-0.90231)+\ln(1-0.04611)] = 15.42556 > \chi^2_{pq;\alpha} = \chi^2_{4;0.05} = 9.4877.\end{aligned}$$

- Aproximación del estadístico Traza de Pillai-Bartlett:

$$\begin{aligned}V_1 =& \lambda_1^2+\lambda_2^2 = 0.90231+0.04611 = 0.94842,\\ F_{V_1} =& \frac{(n-p-q+m-1)V_1}{(|p-q|+m)(m-V_1)} = \frac{(10-2-2+2-1)(0.94842)}{(0+2)(2-(0.94842))} = \\ =& \frac{7(0.94842)}{2(2-0.94842)} = 3.1566 > F_{m(|p-q|+m),m(n-p-q+m-1)} = F_{4,14;0.05} = 3.06.\end{aligned}$$

- Aproximación del estadístico Traza de Hotelling-Lawley:

$$\begin{aligned}U_1 =& \frac{\lambda_1^2}{1-\lambda_1^2}+\frac{\lambda_2^2}{1-\lambda_2^2} = \frac{0.90231}{1-0.90231}+\frac{0.04611}{1-0.04611} = 9.2848,\\ F_{U_1} =& \frac{[m(n-q-p-2)+2]}{m^2(|q-p|+m)}U_1 = \frac{[2(10-2-2-2)+2]}{2^2(0+2)}9.2848 = \\ =& \frac{10}{8}9.2848 = 11.606 > F_{m(|p-q|+m),m(n-p-q-2)+2} = F_{4,10;0.05} = 3.48.\end{aligned}$$

En todos los casos, la conclusión es que se rechaza H_0, es decir, en base a estos datos se descarta la independencia entre las variables relacionadas con la producción de leche y las relacionadas con la conformación de los individuos bajos estudio. También sabemos, en base al contraste planteado, que al menos la primera de las correlaciones es significativa. Veamos cómo determinar si el resto lo son o no.

4.4. Significación de las correlaciones canónicas

Como ya hemos comentado, otra cuestión que tiene interés estudiar dentro de esta técnica es cuál o cuáles de las correlaciones canónicas que se obtienen son significativas.

En la sección de antes hemos visto que para determinar si existe al menos una correlación significativa se resuelve el contraste:

$$\begin{cases} H_0 : \rho_1 = \ldots = \rho_m = 0 \\ H_1 : \text{algún } \rho_i \neq 0 \ i \geq 1 \end{cases}$$

utilizando para ello cualquiera de los estadísticos indicados. Si se rechaza la hipótesis nula, como ya hemos comentado anteriormente, sabemos que al menos la primera correlación es significativa.

Para comprobar si el resto lo son, el siguiente paso consiste en eliminar el efecto del primer par de componentes canónicos, quedando el siguiente contraste de hipótesis:

$$\begin{cases} H_0 : \rho_2 = \ldots = \rho_m = 0 \\ H_1 : \text{algún } \rho_i \neq 0 \ i \geq 2. \end{cases}$$

Para resolverlo eliminamos λ_1^2 del estadístico Λ de Wilks, obteniendo

$$\Lambda_2 = \prod_{i=2}^{m}(1-\lambda_i^2).$$

Si se rechaza este test, se puede concluir que al menos la segunda correlación es significativamente no nula. Se podría reiterar este procedimiento, eliminando en cada ocasión el efecto de la componente que ya se sabe que tiene correlación significativa, hasta que no se rechace el test. En ese momento se tendría que las correlaciones anteriores son todas significativamente no nulas pero que las que se están contrastando en ese momento ya no lo son. Todo esto gracias al orden que guardan las correlaciones. En el $k-$ésimo paso el estadístico para resolver el test sería:

$$\Lambda_k = \prod_{i=k}^{m}(1-\lambda_i^2)$$

que se distribuye según una $\Lambda_{p-k+1,n-k-q,q-k+1}$. Como ya sabemos, existen aproximaciones para este estadístico:

$$\chi_k^2 = -[n-0.5(p+q+3)]\ln\Lambda_k \overset{n\to\infty}{\underset{H_0}{\rightsquigarrow}} \chi^2_{(q-k+1)(p-k+1)}$$

y

$$F_k = \frac{wt - \frac{1}{2}(p-k+1)(q-k+1)+1}{(p-k+1)(q-k+1)} \frac{1-\Lambda_k^{1/t}}{\Lambda_k^{1/t}} \overset{n\to\infty}{\underset{H_0}{\rightsquigarrow}} F_{(p-k+1)(q-k+1),wt-\frac{1}{2}(p-k+1)(q-k+1)+1}$$

donde ahora $w = n - \frac{1}{2}(p+q+3)$ y $t = \sqrt{\frac{(p-k+1)^2(q-k+1)^2-4}{(p-k+1)^2+(q-k+1)^2-5}}$.

También se podrían usar los otros estadísticos estudiados en la sección anterior.

Ejemplo: Continuando con el ejemplo de antes y siguiendo con la hipótesis de normalidad multivariante, ya que sabemos que al menos una correlación es significativa, veamos si ambas lo son.

$$\begin{cases} H_0 : \rho_2 = 0 \\ H_1 : \rho_2 \neq 0 \end{cases}$$

Para resolver este test podemos usar cualquiera de estos estadísticos, como vimos en la sección anterior:

$$\Lambda_2 = (1-\lambda_2)^{=}(1-0.04611) = 0.95389,$$
$$F_2 = \frac{wt - \frac{1}{2}(p-1)(q-1)+1}{(p-1)(q-1)} \frac{1-\Lambda_2^{1/t}}{\Lambda_2^{1/t}}$$

con $w = n - \frac{1}{2}(p+q+3) = 13/2$ y $t = \sqrt{\frac{(p-1)^2(q-1)^2-4}{(p-1)^2+(q-1)^2-5}} = \sqrt{\frac{-3}{-3}} = 1$. Luego

$$F_2 = 7\,\frac{1-0.95389}{0.95389} = 0.33837 < F_{(p-1)(q-1),wt-\frac{1}{2}(p-1)(q-1)+1;0.05} = F_{1,7;0.05} = 5.59.$$
$$\chi_2^2 = -\,[n-0.5(p+q+3)][\ln \Lambda_2] =$$
$$= -\,[10-0.5\cdot 7][\ln 0.95389] = 1.841073 < \chi^2_{(p-1)(q-1);\alpha} = \chi^2_{1;0.05} = 3.8415$$

En este caso no se rechaza la H_0, por lo tanto la segunda correlación canónica no es significativa y sí lo es la primera, en consecuencia del primer contraste realizado en el ejemplo anterior. En base a estos resultados se puede concluir que existe correlación entre las variables bajo estudio (producción y forma) y son fiables las conclusiones que obtuvimos al analizar las correlaciones de las primeras variables canónicas, ya que esta es significativa, mientras que las conclusiones de las segundas ya no son tan fiables, puesto que su correlación no es significativa.

4.5. Aplicación con R

En esta sección vamos a ver cómo se aplica el análisis de correlación canónica a un conjunto de datos mediante el software estadístico R. Los datos provienen de un estudio sobre el comportamiento electoral en Cataluña incluido en el libro de Cuadras ([2]). Se consideran los resultados de unas elecciones celebradas en las 41 comarcas catalanas, y para cada comarca se tabulan, entre otros, los valores de las siguientes variables:

- X_1 = log(porcentaje de votos a CU),
- X_2 = log(porcentaje de votos a PSC),
- X_3 = log(porcentaje de votos a PP),
- X_4 = log(porcentaje de votos a ERC),
- Y_1 = log(cociente Juan/Joan),
- Y_2 = log(cociente Juana/Joana),

siendo CU (Convergencia y Unión), PP (Partido Popular), PSC (Partido Socialista de Cataluña), ERC (Esquerra Republicana). El cociente Juan/Joan significa el resultado de dividir el número de hombres que se llaman Juan por el número de hombres que se llaman Joan. Valores positivos de las variables Y_1 y Y_2 en una comarca indican predominio de los nombres en castellano sobre los nombres en catalán.

Vamos a realizar un análisis de correlaciones canónicas para estas variables que estudie la relación que hay entre el partido al que se vota y el predominio de los nombres en castellano frente a los nombres en catalán.

Para comenzar leeremos los datos que tenemos almacenados en un archivo con formato `.xlsx` a partir de la función `read_excel` de la librería `readxl`. Después los guardaremos en `datos` y copiaremos a una matriz las columnas del archivo que contienen las variables que centran nuestro interés.

```
library(readxl)
datos <- read_excel("Aplicacion.xlsx")
matriz.datos<-datos[c("X1","X2","X3","X4","Y1","Y2")]
```

A continuación separaremos las variables X de las Y, guardándolas en matrices distintas.

```
datos.Partidos<-datos[c("X1","X2","X3","X4")]
datos.Nombres<-datos[c("Y1","Y2")]
```

Para el análisis de correlación canónica existen en R dos librerías fundamentales, `CCA` ([5]) y `CCP` ([7]). La primera de ellas proporciona herramientas gráficas y de análisis para determinar las correlaciones canónicas de un conjunto de datos, mientras que la segunda incluye los test para resolver la significación de las correlaciones canónicas determinadas.

4.5.1. CCA

En la librería `CCA`, como ya hemos comentado, existen diversas funciones que nos van a servir para revisar la correlación entre las variables, tanto gráfica como analíticamente, y para determinar las correlaciones canónicas, entre otras.

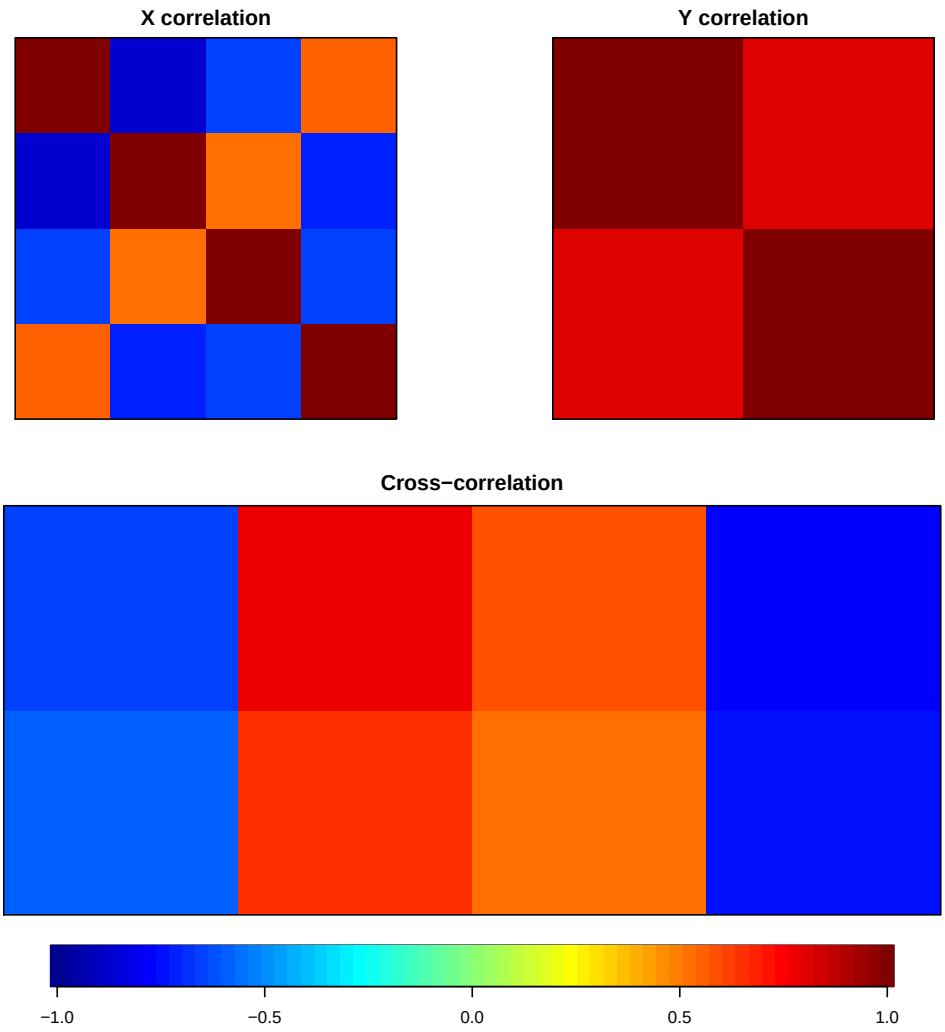
Para determinar las correlaciones entre las variables bajo estudio se puede usar la función `matcor`

```
library(CCA)
matriz.correlacion<-matcor(datos.Partidos,datos.Nombres)
matriz.correlacion

## $Xcor
##            X1         X2         X3         X4
## X1  1.0000000 -0.8520641 -0.6536002  0.5478771
## X2 -0.8520641  1.0000000  0.5127278 -0.7101595
## X3 -0.6536002  0.5127278  1.0000000 -0.6265363
## X4  0.5478771 -0.7101595 -0.6265363  1.0000000
##
## $Ycor
##           Y1        Y2
## Y1 1.0000000 0.8027608
## Y2 0.8027608 1.0000000
```

```
##
## $XYcor
##            X1         X2         X3         X4         Y1         Y2
## X1  1.0000000 -0.8520641 -0.6536002  0.5478771 -0.6403680 -0.5906932
## X2 -0.8520641  1.0000000  0.5127278 -0.7101595  0.7555523  0.6393431
## X3 -0.6536002  0.5127278  1.0000000 -0.6265363  0.5911714  0.5146228
## X4  0.5478771 -0.7101595 -0.6265363  1.0000000 -0.7528501 -0.7448047
## Y1 -0.6403680  0.7555523  0.5911714 -0.7528501  1.0000000  0.8027608
## Y2 -0.5906932  0.6393431  0.5146228 -0.7448047  0.8027608  1.0000000
```

la cual, como podemos ver, nos proporciona la matriz de correlaciones del vector X, del vector Y y la de X con Y. También podemos conseguir una representación gráfica de las correlaciones usando la función `img.matcor`

```
img.matcor(matriz.correlacion, type = 2)
```



Ya sea mediante las correlaciones o mediante su representación gráfica, es evidente que existe una fuerte correlación entre las componentes del vector X, entre las del vector Y y entre las componentes de X e Y, hipótesis de partida de esta técnica.

Una vez comprobado que tiene sentido el análisis deseado, se obtienen las correlaciones canónicas mediante la función `cc`. Para obtener la primera correlación canónica se selecciona la primera componente del vector `cor` que devuelve la función `cc`.

```
cca.datos<-cc(datos.Partidos,datos.Nombres)
cca.datos$cor[1]

## [1] 0.8377282
```

En este caso la primera correlación canónica es 0.8377282 y para ver las primeras variables canónicas se aplica

```
cca.datos$xcoef[,1]

##         X1         X2         X3         X4
## -0.5890329  1.5524428  0.3281909 -1.4678657

cca.datos$ycoef[,1]

##        Y1        Y2
## 0.9632344 0.4594387
```

de donde se deduce que $U_1 = -0.5890329X_1 + 1.5524428X_2 + 0.3281909X_3 - 1.4678657X_4$ y que $V_1 = 0.9632344Y_1 + 0.4594387Y_2$.

Para interpretar las variables canónicas obtenidas, como se vio anteriormente, se pueden obtener las correlaciones entre las variables originales y su correspondiente canónica. Para ello se usa

```
cca.datos$scores$corr.X.xscores[,1]

##         X1         X2         X3         X4
## -0.7796372  0.8965160  0.7072965 -0.9370009

cca.datos$scores$corr.Y.yscores[,1]

##        Y1        Y2
## 0.9792698 0.9069063
```

En vista de estos datos es evidente que U_1 representa la proporción de votos nacionalistas frente a independentistas, ya que las variables que representan a los partidos nacionalistas tienen correlación positiva con esta nueva variable mientras que las que representan a los partidos independentistas la tienen negativa, con una mayor correlación positiva con el partido PSC y

negativa con ERC. Por otro lado se puede ver como V_1 está fuertemente correlacionado con las dos variables de Y, ligeramente más con Y_1 que como sabemos representa el predominio de los nombre castellanos en los hombres, por lo tanto esta variable canónica será mayor si en la comarca hay más proporción de nombres castellanos que catalanes.

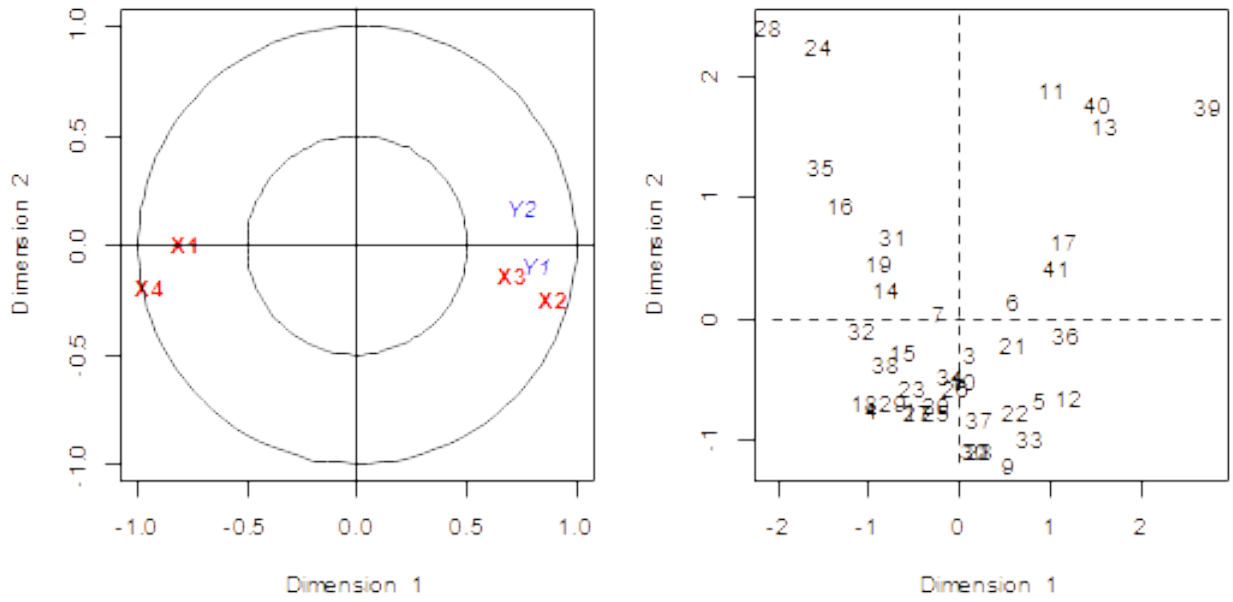
Con respecto a la segunda correlación canónica y su correspondiente variable, se puede repetir el mismo procedimiento

```
cca.datos$cor[2]

## [1] 0.4125206

cca.datos$xcoef[,2]

##         X1         X2         X3         X4
## -13.588911 -10.012398  -3.256534  -4.086477

cca.datos$ycoef[,2]

##        Y1        Y2
## -2.073336  2.221139

cca.datos$scores$corr.X.xscores[,2]

##           X1           X2           X3           X4
##  0.009373661 -0.240361701 -0.130830537 -0.189442382

cca.datos$scores$corr.Y.yscores[,2]

##         Y1         Y2
## -0.2025602  0.4213323
```

Como ya sabíamos esta correlación canónica es menor que la primera y, en este caso, la segunda variable canónica obtenida a partir del vector X, U_2, es difícilmente interpretable, ya que las correlaciones con todas las variables son muy bajas, mientras que la V_2 muestra una clara mayor correlación con la proporción de votos de las mujeres.

Finalmente, se pueden obtener un par más de gráficos que pueden ayudar a la interpretación de las correlaciones canónicas.

```
plt.cc(cca.datos,var.label=TRUE)
```

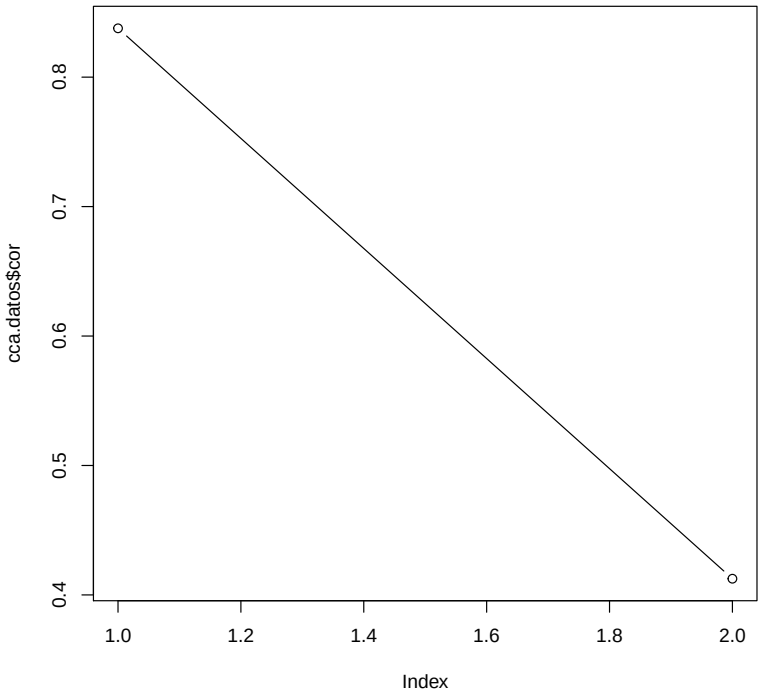


El primero de ellos nos muestra gráficamente los valores de las correlaciones entre las componentes canónicas obtenidas y las variables bajo estudio de las dos primeras componentes, indicando cada componente en una dimensión: (R_{X,U_1}, R_{X,U_2}) y (R_{Y,V_1}, R_{Y,V_2}). El segundo muestra las puntuaciones de los individuos en las dos nuevas componentes.

También se puede obtener un gráfico de las correlaciones canónicas obtenidas, el cual puede ayudar a determinar cuáles son significativas. Este gráfico tiene más sentido cuando el conjunto de variables bajo estudio es mayor y, por lo tanto, se obtienen más correlaciones canónicas, ya que se puede apreciar como estas van decreciendo. En nuestro caso no es un gráfico que aporte mucho.

De todas formas, la mejor forma de analizar si las correlaciones son significativas o no, como se ha estudiado previamente, es realizando los correspondientes tests.

```
plot(cca.datos$cor,type="b")
```



4.5.2. CCP

La librería `CCP` contiene, entre otras, funciones para aplicar los tests que resuelven los contrastes de significación de las correlaciones estudiados. En particular vamos a usar al función `p.asym` de dicha librería. Esta función requiere los siguientes datos:

- **rho**: vector con las correlaciones canónicas.
- **N**: número de observaciones/individuos.
- **p**: número de variables independientes, es decir la dimensión de X.
- **q**: número de variables dependientes, es decir la dimensión de Y.
- **tstat**: el test que se quiere usar. Por defecto, si no se indica nada, se utiliza el de "Wilks". También se pueden usar el de "Hotelling" y el de "Pillai". Para todos los tests calcula el estadístico original y su aproximación con distribución F. Además, la función resuelve m tests: el primero incluyendo todas las correlaciones canónicas $H_0 : \rho_1 = \rho_2 = \ldots = \rho_m = 0$, el segundo eliminando ρ_1, y así sucesivamente hasta que sólo quede $H_0 : \rho_m = 0$.

Si la usamos en nuestro ejemplo con el test que tiene designado por defecto, se obtiene lo siguiente:

```
library(CCP)
rho <- cca.datos$cor
n <- dim(datos.Partidos)[1]
p <- length(datos.Partidos)
q <- length(datos.Nombres)
p.asym(rho, n, p, q)

## Wilks' Lambda, using F-approximation (Rao's F):
##               stat   approx df1 df2      p.value
## 1 to 2:  0.2474638 8.839448   8  70 2.833654e-08
## 2 to 2:  0.8298268 2.460849   3  36 7.834044e-02
```

Como puede verse, en este caso el test se ha aplicado dos veces. El primero resuelve el contraste de $H_0 : \rho_1 = \rho_2 = 0$ y el segundo el de $H_0 : \rho_2 = 0$. En el primer contraste el p-valor$<$0.05 pero en el segundo dicho valor es p-valor$=$0.07834$>$0.05. Por lo tanto, si somos estrictos, a un nivel de significación del 5 %, sólo la primera correlación canónica es significativa y, por lo tanto, sólo la interpretación de ella sería fiable.

Si usamos cualquiera de los otros test estudiados, el resultado es el mismo.

```
p.asym(rho, n, p, q, tstat="Hotelling")

##  Hotelling-Lawley Trace, using F-approximation:
##               stat    approx df1 df2      p.value
## 1 to 2:  2.5583967 10.873186   8  68 1.006169e-09
## 2 to 2:  0.2050708  2.460849   3  72 6.950754e-02

p.asym(rho, n, p, q, tstat="Pillai")

##  Pillai-Bartlett Trace, using F-approximation:
##               stat   approx df1 df2      p.value
## 1 to 2:  0.8719618 6.956907   8  72 9.377537e-07
## 2 to 2:  0.1701732 2.355991   3  76 7.847480e-02
```

Bibliografía

[1] Cooley, W.W. and Lohnes, P. R. (1971) Multivariate Data Analysis. Wiley, N. York.

[2] Cuadras, C. M. (2012) Nuevos métodos de Análisis Multivariante. CMC Editions, Barcelona.

[3] Cuadras, C. M. and M. Sánchez-Turet (1975) Aplicaciones del análisis multivariante canónico en la investigación psicológica. *R*ev. Psicol. Gen. Aplic., 30, 371-382.

[4] Gittings, R. (1985) Canonical Analysis. A Review with Applications in Ecology. Springer-Verlag, Berlin.

[5] González, I., Déjean, S., Martin, P. G. P. y Baccini, A. (2008) CCA: An R Package to Extend Canonical Correlation Analysis. *J*ournal of Statistical Softward, 23, 12.

[6] Hotelling, H. (1936) Relations between two sets of variates. *B*iometrika, 28(3/4), 321-377.

[7] Menzel, U. (2015) Significance Tests for Canonical Correlation Analysis (CCA). https://cran.r-project.org/web/packages/CCP/CCP.pdf

[8] Muirhead, R. J. (1982) Aspects of Multivariate Statistical Theory. Wiley, N. York.

[9] Rencher, A. C. (1995) Methods of Multivariate Analysis. Wiley, N. York.