Título: Acerca de la detección de la colinealidad en los modelos de regresión lineal.

Autores:

Antonio S. Andújar Rodríguez
Departamento de Estadística y Matemática Aplicada
Universidad de Almería

José García Pérez
Departamento de Economía Aplicada
Universidad de Almería

José Soto Liria
Catedrático de Matemáticas
I.E.S. "Nicolás Salmerón"

Resumen:

En la mayoría de la literatura sobre detección de la colinealidad en el modelo de regresión lineal, se acepta como herramienta el número de condición. Nosotros planteamos el problema buscando la mínima distancia entre cada variable del modelo y el subespacio generado por las restantes ("número métrico"). Empíricamente, hemos calculado el número métrico y de condición de matrices generadas aleatoriamente, concluyendo que las distribuciones de ambos varían con la dimensión de la matriz de datos, lo que no se ajusta a los supuestos de Belsey. Para cada dimensión de la matriz, puede establecerse un número métrico crítico que, al 95%, 99%, etc. de confianza, nos detecte el grado de cuasi-colinealidad en matrices generadas aleatoriamente.

I. INTRODUCCIÓN.

Consideremos el modelo lineal ordinario

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (1)$$

donde $\mathbf{X}$ es una matriz de dimensión $n \times p$, en la cual se recogen p muestras de n datos observados, tomada cada una de ellas de entre p variables explicativas, que están representadas por las columnas de $\mathbf{X}$. Desde el punto de vista geométrico, cada vector columna de $\mathbf{X}$ puede representarse por un punto en el espacio real *n*-dimensional, $\mathbb{R}^n$.

Los p vectores columna de $\mathbf{X}$, que denotaremos por $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\}$, son generadores de un subespacio de $\mathbb{R}^n$, que denotamos $\prec_p$. El rango de la matriz $\mathbf{X}$ determina la dimensión del subespacio. Por otra parte, $\boldsymbol{\beta}$ es un vector de p parámetros y $\boldsymbol{\varepsilon}$ un vector *n*-dimensional de perturbaciones aleatorias de media $E(\boldsymbol{\varepsilon}) = 0$ y varianza $\mathrm{var}(\boldsymbol{\varepsilon}) = \boldsymbol{\sigma}^2\mathbf{I}$, siendo $\boldsymbol{\sigma}^2$ la varianza de la población.

El vector $\mathbf{y}$, de la variable respuesta, también pertenece a este espacio *n*-dimensional. Su esperanza matemática, $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, es una combinación lineal de las columnas de $\mathbf{X}$, esto es, $E(\mathbf{y})$ pertenece al subespacio $\prec_p$. El punto exacto donde está $E(\mathbf{y})$ depende de los verdaderos y desconocidos parámetros β_i, componentes de $\boldsymbol{\beta}$. El vector $\mathbf{y}$ está determinado por un punto, cuya proximidad a $E(\mathbf{y})$ viene dada por el término aleatorio $\boldsymbol{\varepsilon}$ -Véase Figura 1-.

El modelo (1) establece que el vector $\mathbf{y}$ es la suma de los vectores $\mathbf{X}\boldsymbol{\beta}$ y $\boldsymbol{\varepsilon}$; mientras que $\mathbf{X}\boldsymbol{\beta}$ pertenece al subespacio $\prec_p$, ni $\boldsymbol{\varepsilon}$ ni $\mathbf{y}$ pertenecen a dicho subespacio.

Si notamos mediante $\hat{\boldsymbol{\beta}}$ los parámetros estimados, que hacen mínimo el módulo del vector $\boldsymbol{\varepsilon}$, se tiene que $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ es una combinación lineal de las columnas de $\mathbf{X}$, $\hat{\mathbf{y}}$ pertenece al subespacio $\prec_p$ y, casi con toda probabilidad, no coincidirá con $E(\mathbf{y})$.

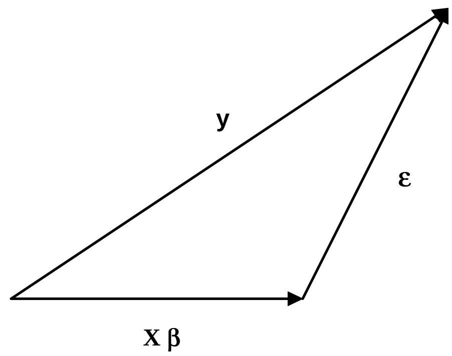

Figura 1

De este modo se tiene –Véase Fig. 2- que $\mathbf{y} = \hat{\mathbf{y}} + \hat{\boldsymbol{\varepsilon}}$, y , además, $\|\boldsymbol{\varepsilon}\| = \|\mathbf{y} - \hat{\mathbf{y}}\|$ es la mínima distancia entre los vectores $\mathbf{y}$ e $\hat{\mathbf{y}}$, es decir, $\hat{\mathbf{y}}$ es la proyección ortogonal de $\mathbf{y}$ sobre el subespacio $\prec_p$. Por tanto, $\hat{\boldsymbol{\varepsilon}}$ es ortogonal a $\prec_p$, verificándose

$$\mathbf{x}_i' \hat{\boldsymbol{\varepsilon}} = 0, \quad i = 1, 2, \ldots, p \qquad (2)$$

siendo ésta, además, una caracterización del vector de errores $\hat{\boldsymbol{\varepsilon}}$, o equivalentemente del vector proyección $\hat{\mathbf{y}}$.

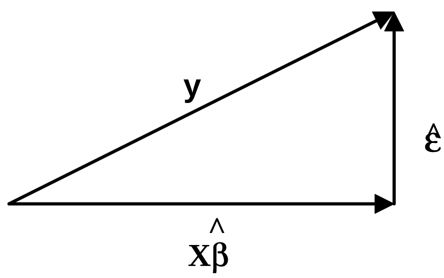

Figura 2

II. COLINEALIDAD Y CUASI-COLINEALIDAD.

La existencia de ***colinealidad*** entre los regresores -variables explicativas-, la interpretamos en el sentido de que exista alguna combinación lineal del tipo

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \ldots + \lambda_p \mathbf{x}_p = 0 \qquad (3)$$

con algún λ_i no nulo.

La colinealidad, o existencia de una dependencia lineal entre las columnas de $\mathbf{X}$, tiene el inconveniente grave de que, bajo su existencia, no se puede estimar el valor de los parámetros $\boldsymbol{\beta}$, al no ser inversible la matriz $\mathbf{X}'\mathbf{X}$. Con datos obtenidos experimentalmente, la aparición de colinealidad perfecta entre los regresores es prácticamente imposible. Por este motivo, se distingue en la literatura al uso, entre colinealidad (ó colinealidad perfecta) y lo que hemos denominado en este trabajo "*cuasi-colinealidad*".

Definimos la ***cuasi-colinealidad*** entre las variables explicativas $\mathbf{x}_i$, $i = 1,2,\ldots, p$ como la existencia de alguna relación del tipo

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \ldots + \lambda_p \mathbf{x}_p + \boldsymbol{\delta} = 0 \qquad (4)$$

con algún λ_i no nulo y, siendo $\boldsymbol{\delta}$ un vector de errores cuyas componentes son relativamente pequeñas en valor absoluto. En este caso, al contrario de lo que ocurre cuando existe colinealidad perfecta, sí es posible estimar $\boldsymbol{\beta}$, pero los valores estimados son poco fiables.

Por lo general, la situación planteada en (4) no se verifica exactamente, como se expuso anteriormente. Poder determinar, a partir de datos experimentales, una colinealidad perfecta en el sentido anterior es, en la práctica, casi imposible pues, aún en el caso de que ésta existiera de hecho, los errores aleatorios, de medida, de redondeo, etc., harían inviable la obtención exacta de los coeficientes λ_i. Para solventar esta dificultad, se dice que un vector $\mathbf{v}_i$ es ***cuasi-combinación lineal*** de $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_p\}$ si existen ciertas constantes reales, λ_i, verificándose:

$$\mathbf{v}_i \approx \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_{i-1} \mathbf{v}_{i-1} + \lambda_{i+1} \mathbf{v}_{i+1} + \ldots + \lambda_p \mathbf{v}_p \tag{5}$$

Considerando el subespacio vectorial $\mathbf{H}_i$ de $\mathbb{R}^p$, generado por $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_p\}$, se tiene que cierto vector $\mathbf{u}_i \in \mathbf{H}_i$ es la mejor aproximación a $\mathbf{x}_i$ cuando la distancia entre $\mathbf{u}_i$ y $\mathbf{v}_i$ es la mínima entre todas las distancias de $\mathbf{u}_i$ a vectores de $\mathbf{H}_i$, esto es,

$$d(\mathbf{x}_i, \mathbf{u}_i) = \min_{u \in H_i} d(\mathbf{x}_i, \mathbf{u}) = \min_{u \in H_i} \|\mathbf{x}_i - \mathbf{u}\| \tag{6}$$

considerando que se está trabajando en el espacio euclídeo real n-dimensional.

También es factible trabajar, en general, en un espacio normado de dimensión finita, utilizando cualquier otra norma distinta de la euclídea. La diferencia al utilizar una u otra es el objeto a optimizar.

III. ÍNDICES MÉTRICOS Y NÚMERO MÉTRICO.

A la distancia indicada en (6), vamos a denominarla ***Índice métrico asociado a la columna i-ésima***, y se denotará por IM_i**.** Utilizando la caracterización (2), es decir, haciendo uso de métodos clásicos de aproximación en espacios vectoriales dotados de producto interno, hemos calculado la distancia indicada en (6), obteniendo para ella una expresión analítica que viene dada por:

$$IM_i = \begin{cases} 0 & \text{si } \det(\mathbf{X}'\mathbf{X}) = 0 \\ \sqrt{\dfrac{\det(\mathbf{X}'\mathbf{X})}{\mathbf{M}_{ii}}} & \text{si } \det(\mathbf{X}'\mathbf{X}) \neq 0 \end{cases} \tag{7}$$

donde $\mathbf{M}_{ii}$ es el menor de orden $p-1$ de la matriz $\mathbf{X}'\mathbf{X}$ correspondiente al elemento (i,i) de su diagonal principal.

Evidentemente, si para algún i, se tiene que $IM_i = 0$**,** entonces la dependencia lineal es perfecta; $\mathbf{x}_i$ es colineal con $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_p\}$. A medida que los valores IM_i se alejan de cero, la colinealidad disminuye.

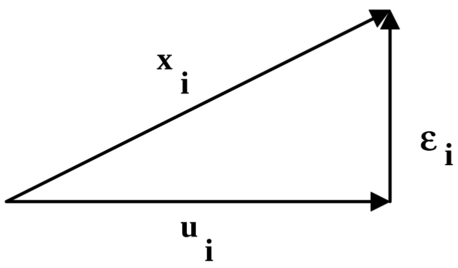

Figura 3

Observando la figura 3, se sigue que:

$$IM_i = d(\mathbf{x}_i, \mathbf{u}_i) = \|\boldsymbol{\varepsilon}_i\| \Rightarrow \operatorname{sen}(\mathbf{x}_i, \mathbf{u}_i) = \frac{IM_i}{\|\mathbf{x}_i\|} \tag{8}$$

Si las columnas de la matriz $\mathbf{X}$ están normalizadas a módulo 1, los índices métricos representan los senos de los ángulos formados por cada vector columna y su mejor aproximación en el subespacio generado por los restantes $p-1$ vectores columna.

En este trabajo tratamos la colinealidad desde un punto de vista geométrico, de modo que cuanto menor sea el índice métrico asociado a un vector columna de la matriz, diremos que es mayor la cuasi-colinealidad.

Definimos ***Número métrico*** asociado a la matriz **X**, al mínimo de los índices métricos de sus vectores columna, y lo denotamos por $NM(\mathbf{X})$, es decir,

$$NM(\mathbf{X}) = \min_{\mathrm{i}}\{IM_i\}. \qquad (9)$$

De este momo se detecta, por una parte, qué columna es la más proxima al subespacio generado por las demás y qué "grado" de proximidad hay. A partir de (9) se puede determinar un valor crítico, ξ, tal que para valores de $NM(\mathbf{X})$ menores que ξ, se aconseje tomar medidas adicionales para corregir el efecto dañino de la colinealidad en la estimación de los parámetros del modelo. El tipo de medida a tomar será objeto de un estudio posterior al que nos ocupa.

El objetivo, en el planteamiento del sistema (1), es encontrar la mejor aproximación al vector $\mathbf{y}$ en el subespacio generado por $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\}$, es decir, encontrar un vector $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ que haga mínima la distancia $d(\mathbf{y}, \hat{\mathbf{y}})$. Siguiendo la definición anterior, podemos considerar el índice IM_y como una medida del error que cometemos al tomar $\hat{\mathbf{y}}$ como valor aproximado de $\mathbf{y}$.

Si el vector $\mathbf{y}$ es unitario, IM_y coincide con el seno del ángulo que forman los vectores $\hat{\mathbf{y}}$ e $\mathbf{y}$, ya que

$$\mathrm{sen}(\hat{\mathbf{y}}, \mathbf{y}) = \frac{IM_y}{\|\mathbf{y}\|} = \frac{IM_y}{1} = IM_y . \qquad (10)$$

De este modo podemos trabajar con los índices métricos o, si se desea, con los ángulos asociados a ellos.

Tipificación de los datos.

Matrices de datos que difieren únicamente en la escala en la que están expresados los mismos, es decir, difieren sólo en las unidades de medida asignadas a las columnas de datos –variables explicativas-, presentan modelos de estructuras esencialmente equivalentes. Sin embargo, tanto entre los índices de condición como entre los índices métricos de dos matrices cuya única diferencia sea la escala en la que están expresados los datos de uno o varios de sus vectores columna, puede haber "grandes" variaciones. De ahí la necesidad de tipificar, por columnas, las matrices de datos correspondientes a modelos estructurales equivalentes, de modo que hagan posible y den un sentido adecuado, tanto a la definición como a la comparación de los índices de condición e índices métricos.

En la mayoría de la literatura sobre regresión cresta, se acepta como proceso normal de estandarización que los datos estén tipificados, de modo que los vectores tengan de módulo uno. Esto produce que en las matrices ortogonales el número de condición tome el valor mínimo que es uno y el número métrico su valor máximo, uno, como corresponde a la ausencia de colinealidad. Cualquier otra forma de tipificar podría no producir estos efectos beneficiosos.

Por otra parte, los cambios de escala en las columnas de la matriz $\mathbf{X}$ no afectan a las combinaciones lineales exactas que pudieran existir entre ellas.

En cuanto a la necesidad de centrar los datos, por lo que respecta al número métrico, ésta surge cuando queremos comparar datos que varían en rangos distintos. Por ejemplo:

$$x_{ik} \in [-10,10] \text{ y } x_{jk} \in [20,50],\ i \neq j\ ,\ \ k = 1,2,\ldots,n$$

no son comparables, salvo que estén centrados respecto a la media y normalizados a módulo uno, en cuyo caso ambos variarían en el intervalo $[-1,+1]$. En todo caso, esto es aceptado en la mayoría de la literatura sobre el tema de la colinealidad, en que una de las hipótesis empleada es que $\mathbf{X}'\mathbf{X}$ esté en forma de matriz de correlación, lo cual lleva implícita la hipótesis de que los datos estén tipificados.

El número de condición.

En la literatura sobre regresión lineal, es aceptado ampliamente como una de las mejores herramientas para la detección de colinealidad el *número de condición*, $k(\mathbf{X})$, que viene definido para una matriz de datos $\mathbf{X}$ como

$$k(\mathbf{X}) = \frac{\mu_{\text{máx}}}{\mu_{\text{mín}}} \tag{11}$$

siendo μ_i, $i = 1,2,\ldots,p$, los valores singulares asociados a la matriz $\mathbf{X}$ en la descomposición singular:

$$\mathbf{X} = \mathbf{UDV}',\ \text{con}\ \mathbf{U}'\mathbf{U} = \mathbf{I},\ \mathbf{V}'\mathbf{V} = \mathbf{I},\ \mathbf{D} = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_p)$$

Según Belsey et *al*. (1980), para el número de condición se considera la serie 10, 30, 90, 300, 900, etc., correspondientes a valores de R^2 dados por 0.9, 0.99, 0.999, etc., de modo que, para $k(\mathbf{X}) \leq 10$ existe colinealidad débil; si $10 < k(\mathbf{X}) < 30$, la colinealidad es moderada; y si $k(\mathbf{X}) \geq 30$, la colinealidad es fuerte.

Para llevar a cabo sus experimentos, Belsey *et al.* señalan:

"*El número de observaciones, n, que no es importante, es alrededor de 24-27, y 3-5 variables, dependiendo del experimento*".

A nosotros nos parece que esta afirmación no es del todo correcta. La colinealidad depende de la dimensión de la matriz y, por consiguiente, tanto del número de observaciones como del número de variables, como hemos comprobado experimentalmente generando matrices aleatorias, calculando sus números de condición y estudiando su comportamiento. Este asunto se tratará en los siguientes apartados.

IV. ESTUDIO EXPERIMENTAL DE LA VARIACIÓN DE LOS NÚMERO DE CONDICIÓN Y MÉTRICO, CON LA DIMENSIÓN DE LA MATRIZ DE DATOS.

Se han generado 100 matrices aleatorias con elementos en el rango $[-1,+1]$, de forma que el tamaño de las muestras es siempre superior al número de variables explicativas. Una vez tipificadas, hemos calculado sus números de condición y se han distribuido agrupándolos en cuatro clases, correspondiendo a los intervalos $[0,10),[10,30),[30,90),[90,+\infty)$, que corresponden a cuatro grados posibles de cuasi-colinealidad: inexistencia, débil, fuerte y muy fuerte, respectivamente. Los resultados se han resumido en la tabla 1.

Como puede observarse, el número de condición depende tanto del número de variables como del tamaño muestral.

Cuanto mayor es la diferencia entre el tamaño muestral y el número de variables, menor es la probabilidad de existencia de colinealidad, para estas matrices generadas aleatoriamente. De hecho, a partir de cierta dimensión, todas las matrices tienen su número condición menor de 10, lo que indica ausencia de cuasi-colinealidad. (obsérvense, por ejemplo, la filas 4,5 y 10 de la tabla 1).

Tamaño Muestral	Número Variables	NC[1] <10	10<NC<30	30<NC<90	NC>90
4	3	68	20	7	5
5	3	91	8	0	1
6	3	98	2	0	0
7	3	100	0	0	0
8	3	100	0	0	0
5	4	52	29	10	9
6	4	85	14	0	1
7	4	92	8	0	0
8	4	98	2	0	0
9	4	100	0	0	0

Tabla 1

Hemos realizado un estudio similar con el número métrico. En este caso se han generado tres experiencias con 1000 matrices aleatorias, de las cuales se ha calculado su número métrico, habiendo tipificado los datos por columnas previamente.

En la figura 4, se ha representado en color rojo la distribución del número métrico de 1000 matrices aleatorias tipificadas de dimensión 5×3, de forma que en el eje de abcisas se colocan los posibles valores del número métrico y en el de ordenadas, la frecuencia absoluta de ellos. Análogamente, en color verde se ha representado la distribución de 1000 matrices 5×4 y, en color azul, la de 1000 matrices 7×4.

Comparando las distribuciones de números métricos de las matrices 5×3 y de las matrices 5×4, se observa que para muestras del mismo tamaño, el incremento del número de variables produce que haya más probabilidad de que la matriz obtenida tenga cuasi-colinealidad en sus columnas.

Análogamente, si se comparan las distribuciones de las matrices 5×4 y de las de dimensión 7×4, se observa que al aumentar el tamaño de la muestra manteniendo fijo el número de variables explicativas, se produce que el número de matrices con cuasi-colinealidad, en las condiciones indicadas, disminuye.

[1] NC siglas de número condición

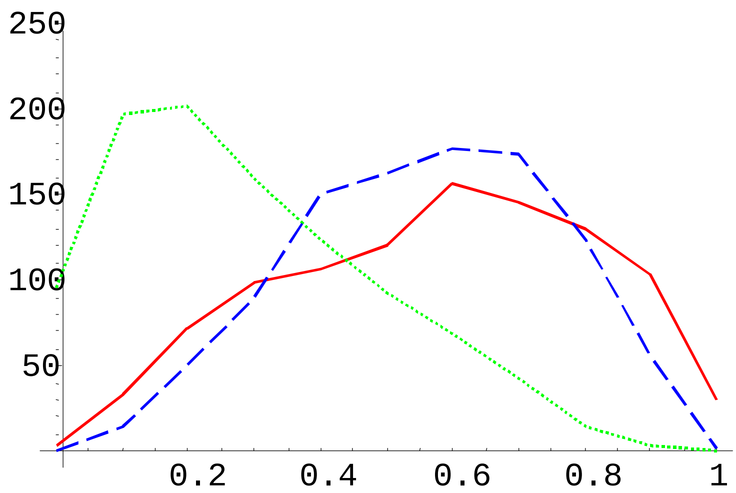

Figura 4: Distribuciones de NM

Este comportamiento es general, existiendo grandes variaciones cuando los tamaños muestrales y el número de variables son cantidades próximas. Por otra parte, al aumentar la diferencia entre el tamaño de la muestra y el número de variables, disminuye drásticamente la probabilidad de que una matriz aleatoria tenga cuasi-colinealidad.

VI. RELACIÓN ENTRE EL NÚMERO CONDICIÓN Y EL NÚMERO MÉTRICO.

En esta sección se desea comprobar si existen similitud y/o divergencias entre las dos medidas que hemos presentado en las secciones precedentes, el número condición y el número métrico que proponemos. Para ello, hemos generado 1000 matrices 7×6 y, una vez tipificadas, se han calculado sus números métrico y condición.

En la figura 5 se han representado en unos ejes coordenados la nube de 1000 puntos (NM, NC) de las matrices indicadas en el párrafo anterior, esto es, la abcisa de cada punto es el número métrico de una matriz, mientras la ordenada representa su número condición.

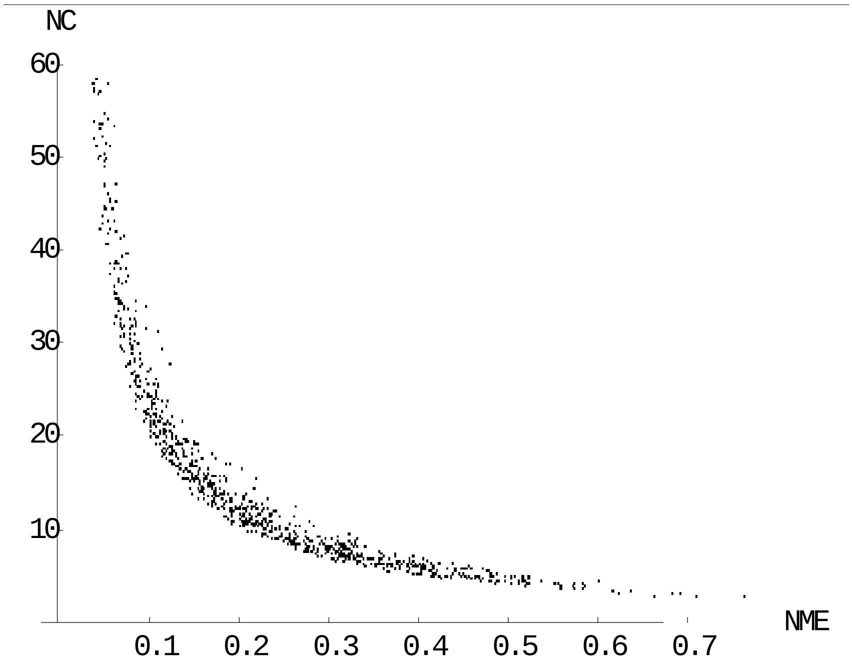

Figura 5: Comparación NM y NC

Por observación directa, dado que estos puntos no se distribuyen por todo el plano, se deduce que hay una cierta relación entre ambos, aunque no es de tipo funcional ($NC = \phi(NM)$). De hecho aparece una "**banda**" de puntos decreciente, de tal modo que al aumentar el número métrico, disminuye el condición y viceversa. Esto es conforme con lo esperado en el sentido de que ambas medidas son una guía de detección de la cuasi-colinealidad y, por el hecho de que el caso de ortogonalidad, es decir, la ausencia total de cuasi-colinealidad, se corresponde con el máximo valor posible del número métrico y con el mínimo valor del número condición (ambos son 1).

Pero, sin embargo, son también notables sus diferencias. En efecto, podemos observar que para cada valor fijo del número condición, el métrico tiene un rango de variación más o menos amplio. Por ejemplo, para $NC \approx 17$, se tiene que se encuetran matrices con $NM \in (0.12,0.20)$, esto es, existen matrices cuyo número métrico se encuentra en una franja de un 8% de variación sobre la variación total del mismo, y mantienen el mismo número condición.

Por otra parte, si dejamos fijo el número métrico, por ejemplo, para $NM \approx 0.1$, se tiene que el número condición varía en un rango de valores entre 20 y 35 aproximadamente. Esto es además muy significativo, dado que estaríamos que mientras una medida -el métrico- está fija, la otra –el condición- puede dar lugar a matrices con cuasi-colinealidad moderada (<30) y fuerte (>30), según Belsey.

Hay que tener en cuenta que el concepto de número condición, a pesar de su amplio uso en este ámbito, no surge para la detección de la cuasi-colinealidad, sino en los problemas de las matrices mal condicionadas y los grandes errores que ello provoca en la resolución de sistemas lineales.

VI. CONCLUSIONES. LÍNEAS DE INVESTIGACIÓN.

El problema de la detección y corrección de la cuasi-colinealidad ha sido abordado en la literatura de diversas formas y se han dado métodos para su detección, pero ninguno ha sido adoptado universalmente.

El más aceptado, quizás, haya sido el del número condición, pero se nos plantean algunas dudas, tal como se ha expuesto en el apartado anterior, acerca de su idoneidad en cualquier caso.

Si se dice que una matriz posee un número de condición 25, 30 ó 100, ¿qué interpretación podemos darle a esto?. El número métrico, sin embargo, es más intuitivo. Se ha deducido desde un punto de vista geométrico, dando el sentido de que cuánto menor sea el ángulo que forma un cierto vector con un hiperplano, más cuasi-colinealidad hay, hasta darse el caso posible de la colinealidad perfecta, que se da cuando el ángulo es de 0º. Valores de este ángulo muy pequeños -lo cual es determinado por el número métrico-, determinan la existencia de fuerte cuasi-colinealidad.

La probabilidad de que cierta matriz tenga multicolinealidad varía con la dimensión de la misma, y para matrices "*desproporcionadas*" en el sentido de poseer muchas más filas que columnas, la probabilidad de que aleatoriamente aparezca cuasi-colinealidad es muy pequeña.

En cuanto a futuras investigaciones, podemos reseñar las siguientes lineas:

- Estudiar la cuasi-colinealidad utilizando otras normas distintas de la eucídea.
- Utilizando los índices métricos, detectar si existen varias cuasi-colinealidades distintas entre las columnas de la matriz $\mathbf{X}$, esto es, dos o más conjuntos disjuntos de vectores columnas que estén en cuasi-combinación lineal.
- Una vez comprobada la existencia de cuasi-colinealidad mediante el número métrico se tratará de intentar solventar el problema de la determinación de los parámetros del modelo lineal, teniendo en cuenta que será conocida qué columna o columnas son las que determinan el índice métrico.

BIBLIOGRAFÍA.

[1] Belsey, D.; Kuh, E.; Welsch, R.E. (1980): *Regression Diagnostics: Identifying Influential Data and Source of Collinearity,* Ed. John Wiley & Sons.

[2] Farrar, D.E., Glauber, R.R. (1967): Multicollinearity Analysis: The Problem Revisited, *Review of Economics and Statistics*, 49, pp. 92-107.

[3] Kincaid, D.; Cheney, W. (1994): *Análisis Numérico*, Ed. Addison-Wesley Iberoamericana.

[4] Mason, R.; Gunst, R.F.; Webster, J.T. (1975): Regression Analysis and Problems of Multicollinearity, *Communications in Statistics*, 4 (3), pp. 277-292.

[5] Silvey, S.D. (1969): Multicollinearity and Imprecise Estimation, *Journal of Royal Statistical Society, Series B*, 31, pp. 539-552.

[6] Wang, S.; Tse, S.; Chow, S. (1990): On the Measures of Multicollinearity in Least Squares Regression, *Statistics and Probability Letter*, 9, pp. 347-355.

[7] Witchers, R. (1975): The Detection of Multicollinearity: A Comment, *The Review of Economics and Statistics,* pp. 366-369.