



SÉRIE DE NOTAS TÉCNICAS EM ECONOMIA DA UFG
NT N. 01

NOTA TÉCNICA EM ECONOMIA
n. 01

Publicação cujo objetivo é auxiliar na elaboração de aulas e pesquisas do Curso de Ciências Econômicas da UFG, divulgando demonstrações técnicas e metodológicas ou rotinas computacionais voltadas para a economia.

As opiniões contidas nesta publicação não representam o ponto de vista do NEPEC ou da FACE/UFG, sendo de inteira responsabilidade de seu(s) autor(es).

Reprodução permitida, desde que citada a fonte.

FACE – Faculdade de Administração, Ciências Contábeis e Ciências Econômicas
Curso de Ciências Econômicas
Direção FACE
Maria do Amparo Albuquerque Aguiar
Vice-Direção FACE
Flávia Regina Cazerneschy
Coordenação do Curso de Ciências Econômicas
Cleyzer Adrian da Cunha

NEPEC – Núcleo de Estudos e Pesquisas Econômicas
Coordenação
Dnilson Carlos Dias

Endereço

Campus Samambaia, Prédio da FACE – Rodovia Goiânia/Nova Veneza, km. 0 – Caixa Postal 131, CEP 74001-970, Goiânia – GO.
Tel. (62) 3521 – 1390

URL

<http://www.face.ufg.br/eco>

Rotina de Cálculo do *Conditional Number* no GRETL

Sandro Eduardo Monsueto
FACE/UFG

NEPEC/FACE/UFG
Goiânia – Setembro de 2010
Versão 1.0



Rotina de calculo do Conditional Number no GRETL

Sandro Eduardo Monsueto¹
Universidade Federal de Goiás

Resumo:

Esta nota apresenta uma rotina para calcular o conditional number, um importante indicador de multicolineariedade, usando o pacote GRETL.

Palavras chave: GRETL, conditional number, multicolineariedade

Abstract:

This note presents a routine to calculate the conditional number, an important multicollinearity indicator, using the package GRETL.

Key words: GRETL, conditional number, multicollinearity

1. Introdução

Esta nota técnica tem por objetivo explicar a construção de uma rotina no software GRETL de análise econométrica para calcular o conditional number, tal como sugerido em Belsley, Kuh, e Welsch (2004). Primeiramente é apresentada a estatística de cálculo do conditional number, bem como alguma de suas características, para depois apresentar a rotina de cálculo no programa. A rotina foi projetada para a versão 1.9.1 de julho de 2010. Recomenda-se uma leitura prévia no manual do programa nas sessões sobre criação e manipulação de matrizes e sobre a utilização de loops.

2. O Conditional Number

O conditional number de uma matriz qualquer é dada pela raiz quadrada da razão entre seu maior autovalor e seu menor autovalor, ou seja:

$$CN = \sqrt{\frac{\text{Maior Autovalor}}{\text{Menor Autovalor}}} \quad (1)$$

Dentro do contexto da análise de regressão, a matriz de interesse é a matriz \mathbf{X} , cujas colunas são formadas pelas variáveis explicativas do modelo estimado. Como tal matriz não é quadrada, é usada a matriz $\mathbf{X}'\mathbf{X}$, que satisfaz à condição de Mínimos Quadrados Ordinários:

¹ monsueto@face.ufg.br

$$X'Xb = X'y \quad (2)$$

Onde X' é a transposta da matriz de variáveis independentes, b é o vetor de coeficientes estimados e y o vetor coluna da variável explicada. Porém, como destacado em Green (2003), os autovalores são sensíveis à escala de medida das variáveis do modelo. Para criar uma padronização, se recomenda o uso de uma matriz “normalizada”, cujas colunas são os vetores da matriz X divididas por sua respectiva norma. A norma de um vetor e é dada por (Greene, 2003):

$$\|e\| = \sqrt{e'e} \quad (3)$$

Portanto, cada coluna da matriz X de dados deverá ser dividida por sua norma para se obter uma matriz P padronizada. Desta forma, o conditional number é calculado sobre uma matriz A , definida por:

$$A = P'P \quad (4)$$

Em síntese, as etapas para a obtenção do conditional number podem ser assim sintetizadas:

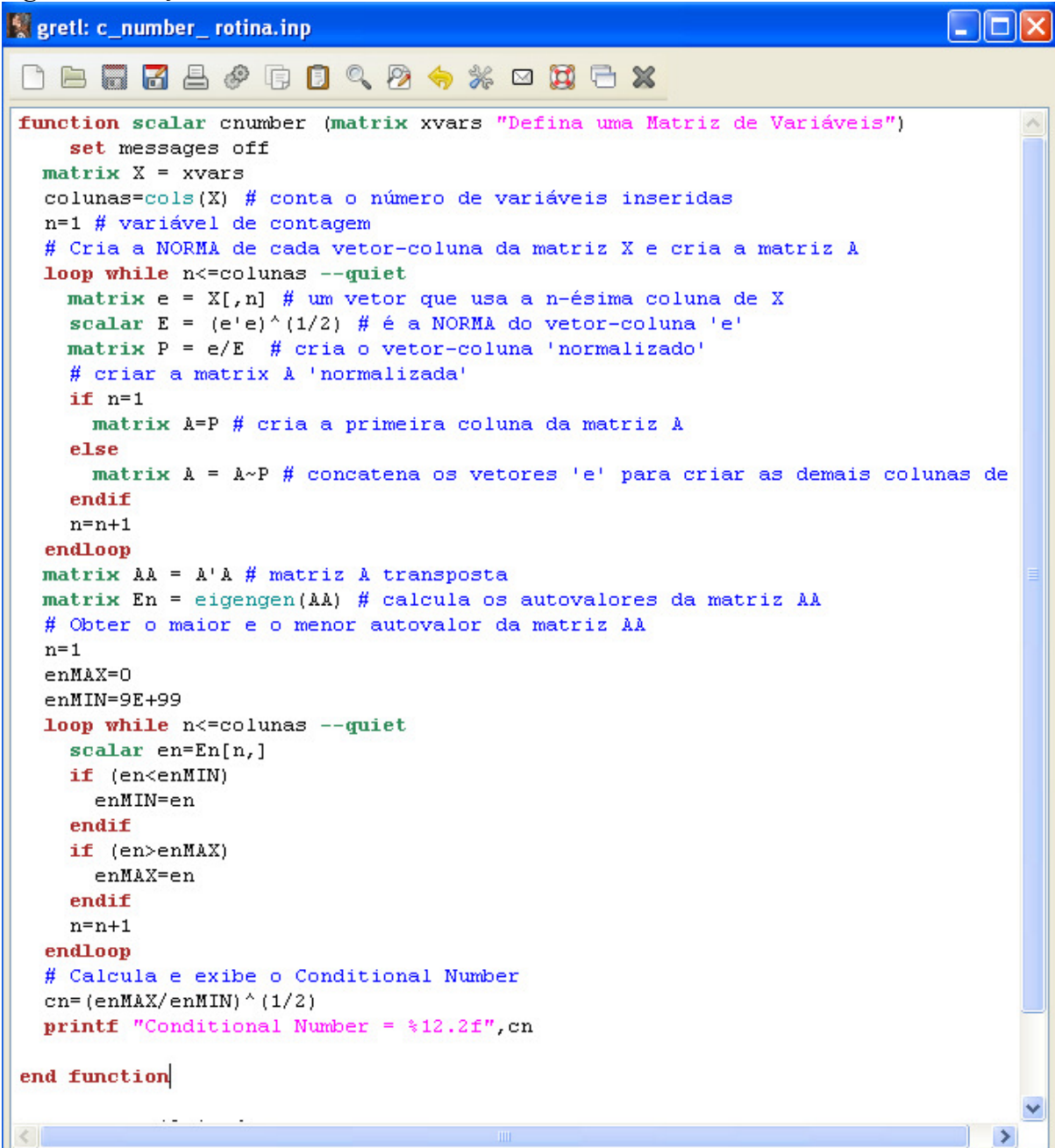
- i. Determinação da matriz X de variáveis explicativas;
- ii. Cálculo da norma de cada coluna da matriz X , com indicado na equação (3);
- iii. Obtenção da matriz P normalizada a partir de cada coluna de X dividida pela sua respectiva norma;
- iv. Obtenção da matriz $A = P'P$
- v. Cálculo do conditional number da matriz A , com base em (1)

Estas etapas são seguidas para a construção da rotina. Segundo Belsley, Kuh, e Welsch (2004) este índice oferece uma medida da sensibilidade do vetor de coeficientes b à pequenas mudanças na base de dados, como exclusão ou inclusão de variáveis e/ou observações. Sendo assim, costuma ser usado como indicador de multicolinearidade entre as variáveis do modelo de regressão. Em geral, um conditional number superior a 20 significa problemas significativos de multicolinearidade.

3. Criando a rotina

A rotina foi criada na forma de uma função chamada *cnumber*, que retorna o conditional number para um conjunto de variáveis informadas. Esta função pode ser instalada no programa pelo menu “Arquivo” → “Arquivo de Funções”, ou simplesmente executada como um arquivo de comandos. A Figura 1 mostra a função como aparece na tela do GRETL e o Quadro 1 mostra a função pronta para ser copiada e editada por outros usuários.

Figura 1. Função *cnumber* – tela do GRETL



```
function scalar cnumber (matrix xvars "Defina uma Matriz de Variáveis")
  set messages off
  matrix X = xvars
  colunas=cols(X) # conta o número de variáveis inseridas
  n=1 # variável de contagem
  # Cria a NORMA de cada vetor-coluna da matriz X e cria a matriz A
  loop while n<=colunas --quiet
    matrix e = X[,n] # um vetor que usa a n-ésima coluna de X
    scalar E = (e'e)^(1/2) # é a NORMA do vetor-coluna 'e'
    matrix P = e/E # cria o vetor-coluna 'normalizado'
    # criar a matrix A 'normalizada'
    if n=1
      matrix A=P # cria a primeira coluna da matrix A
    else
      matrix A = A~P # concatena os vetores 'e' para criar as demais colunas de
    endif
    n=n+1
  endloop
  matrix AA = A'A # matrix A transposta
  matrix En = eigengen(AA) # calcula os autovalores da matrix AA
  # Obter o maior e o menor autovalor da matrix AA
  n=1
  enMAX=0
  enMIN=9E+99
  loop while n<=colunas --quiet
    scalar en=En[n,]
    if (en<enMIN)
      enMIN=en
    endif
    if (en>enMAX)
      enMAX=en
    endif
    n=n+1
  endloop
  # Calcula e exibe o Conditional Number
  cn=(enMAX/enMIN)^(1/2)
  printf "Conditional Number = %12.2f",cn
end function
```

Quadro 1. Função cnumber para editar

```
function scalar cnumber (matrix xvars "Defina uma Matriz de Variáveis")

matrix X = xvars
colunas=cols(X) # conta o número de variáveis inseridas
n=1 # variável de contagem
# Cria a NORMA de cada vetor-coluna da matriz X e cria a matriz A
loop while n<=colunas --quiet
  matrix e = X[,n] # um vetor que usa a n-ésima coluna de X
  scalar E = (e'e)^(1/2) # é a NORMA do vetor-coluna 'e'
  matrix B = e/E # cria o vetor-coluna 'normalizado'
  # criar a matrix A 'normalizada'
  if n=1
    matrix A=B # cria a primeira coluna da matriz A
  else
    matrix A = A~B # concatena os vetores 'e' para criar as demais colunas de A
  endif
  n=n+1
endloop
matrix AA = A'A # matriz A transposta
matrix En = eigengen(AA) # calcula os autovalores da matriz AA
# Obter o maior e o menor autovalor da matriz AA
linhasEn=rows(En) # número de linhas (autovalores) da matriz 'En'
n=1
enMAX=0
enMIN=9E+99
loop while n<=colunas --quiet
  scalar en=En[n,]
  if (en<enMIN)
    enMIN=en
  endif
  if (en>enMAX)
    enMAX=en
  endif
  n=n+1
endloop
# Calcula e exibe o Conditional Number
cn=(enMAX/enMIN)^(1/2)
printf "Conditional Number = %12.2f",cn

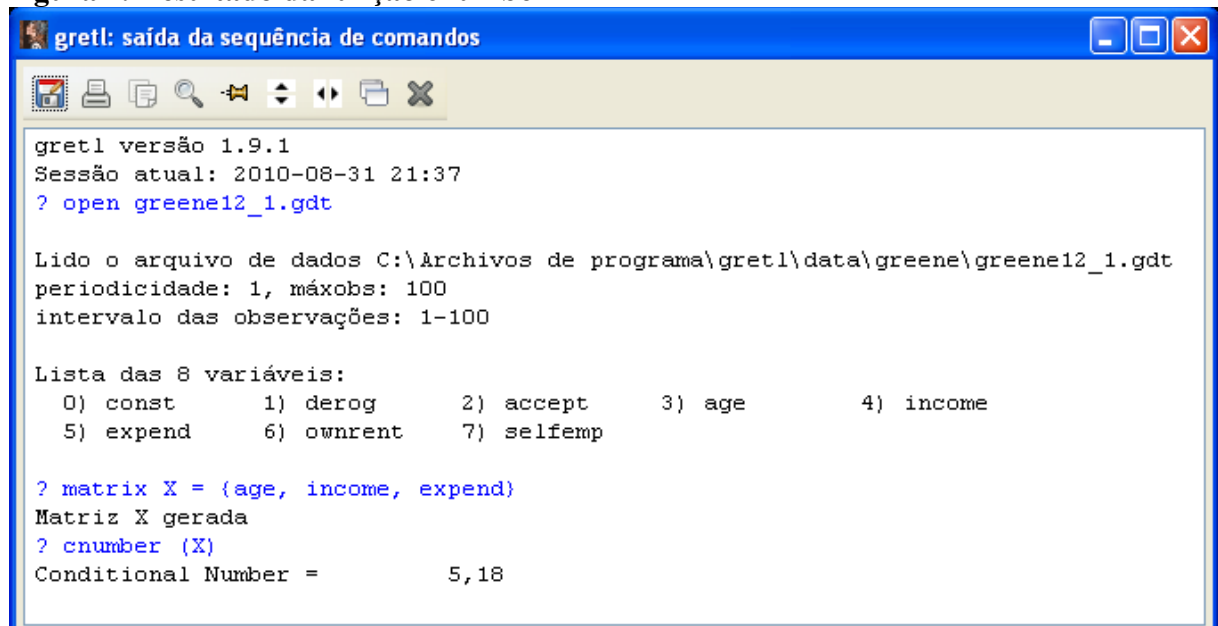
end function
```

Como exemplo de aplicação da função, considere o arquivo *greene12_1.gdt*, disponível na base de dados do GRET. Uma vez que a função já tenha sido executada uma vez, o usuário pode executar a seguinte sequência de comandos:

```
open greene12_1.gdt
matrix X = {age, income, expend}
cnumber (X)
```

A primeira linha abre a base de dados de exemplo, enquanto a segunda linha define uma matrix com os nomes das variáveis usadas. Por fim, a última linha executa a função, obtendo o conditional number. O resultado desta operação pode ser visto na Figura 2.

Figura 1. Resultado da função cnumber



```
gretl: saída da sequência de comandos
gretl versão 1.9.1
Sessão atual: 2010-08-31 21:37
? open greene12_1.gdt

Lido o arquivo de dados C:\Archivos de programa\gretl\data\greene\greene12_1.gdt
periodicidade: 1, máxobs: 100
intervalo das observações: 1-100

Lista das 8 variáveis:
  0) const      1) derog      2) accept      3) age         4) income
  5) expend     6) ownrent    7) selfemp

? matrix X = (age, income, expend)
Matriz X gerada
? cnumber (X)
Conditional Number =          5,18
```

Como mostra o resultado, o conjunto das variáveis gera um conditional number de 5.18, indicando uma baixa colineariedade entre os regressores. Para adicionar a constante à análise, basta acrescentar à matriz X a variável *const*:

$$\text{matrix } X = \{\text{const, age, income, expend}\}$$

Referências bibliográficas

Belsley, D.A.; Kuh, E. e Welsch, R.E. *Regression diagnostics: Identifying influential data and sources of collienarity*, New Jersey: Wiley-IEEE. 2004.

Greene, W.H.. *Econometrics analysis*, New Jersey: Prentice Hall. 5ºed, 2003.