TWO-DIMENSIONAL GAUSS-LEGENDRE QUADRATURE:

SEEMINGLY UNRELATED DISPERSION-FLEXIBLE

COUNT REGRESSIONS

by

Joseph V. Terza
and
Abbie Zhang

Department of Economics
Indiana University School of Liberal Arts at IUPUI
Indianapolis, IN 46202

M-Estimation

-- Consider the following M-Estimator (ME), $\hat{\theta}$, for the parameter vector θ

$$\hat{\theta} = \underset{\breve{\theta}}{\operatorname{argmax}}\, Q_n(\breve{\theta}) \tag{1}$$

where $\underset{\breve{\theta}}{\operatorname{argmax}}$ means "the value of $\breve{\theta}$ at which Q is maximized", $\breve{\theta} \in \Theta$, Θ is the parameter space (i.e. the set of all possible values of θ)

$$Q_n(\breve{\theta}) = \frac{\sum_{i=1}^{n} q(\breve{\theta},\ Z_i)}{n}$$

and $Z_i = [Y_i \quad \mathcal{X}_i]$ = vector of observed data on the outcome (Y_i – possibly multivariate, i.e., a vector of outcomes) and the regressors ($\mathcal{X}_i$), respectively, for the ith observation in a sample of size n (i = 1, ..., n).

-- This ME class subsumes maximum likelihood, nonlinear least squares and method of moments estimators.

M-Estimation and Two-Dimensional (2D) Integration

We focus on ME in which the objective function is of the following form

$$q(\breve{\theta},\ Z_i) = h\left[\int_{a_2}^{b_2}\int_{a_1}^{b_1} q^*(\breve{\theta}, Z_i, \eta_1, \eta_2)\, d\eta_1 d\eta_2\right] \qquad (2)$$

where

-- the a's and the b's are known (or known up to a vector of parameters to be estimated as part of θ)

-- $[\eta_1 \quad \eta_2]$ is typically a bivariate vector of unobservables

and

-- the integral is not closed form.

M-Estimation and 2D Integration (cont'd)

-- Examples that come to mind:

-- Nonlinear models with bivariate endogenous regressors.

-- Nonlinear models with an endogenous regressor and sample selection.

-- Bivariate (Multivariate?) nonlinear seemingly unrelated regressions (SUR)

An Application: Bivariate Dispersion-Flexible Count-Valued SUR

-- We focus here on our proposed variant and extension of the ME of Aitchison and Ho (1989).

-- In our model, the relevant version of (2) has the following form

$$q(\breve{\theta},\ Z_i) = \ln\Big[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{(Y_{1i}\,|\,X_i,\,\eta_1)}(Y_{1i}, X_i, \eta_1) \times f_{(Y_{2i}\,|\,X_i,\,\eta_2)}(Y_{2i}, X_i, \eta_2)\, g(\eta_1, \eta_2)\, d\eta_1 d\eta_2\Big] \qquad (3)$$

where

$f_{(Y_{ji}\,|\,X_i,\,\eta_j)}(Y_{ji}, \mathcal{X}_i, \eta_j) \equiv$ the probability mass function (pmf) of the jth count-valued outcome (Y_{ji}), conditional on the regressors $(\mathcal{X}_i)$ and the relevant unobservable (η_j) for the jth equation.

$g(\eta_1, \eta_2) \equiv$ the joint probability density function (pdf) of the unobservables.

Aitchison, J., & Ho, C. H. (1989): "The Multivariate Poisson-Log-Normal Distribution," *Biometrika*, 76, 643–653. https://doi.org/10.1093/biomet/76.4.643

An Application: Bivariate Dispersion-Flexible Count-Valued SUR (cont'd)

-- The primary motivation for specify the bivariate system of count regressions in this way is to take advantage of possible gains in efficiency (see Zellner, 1962).

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, *57*(298), 348–368.

An Application: Bivariate Dispersion-Flexible Count-Valued SUR (cont'd)

-- Consider the case in which (for j = 1, 2)

$$f_{(Y_{ji} \mid \mathcal{X}_i, \eta_j)}(Y_{ji}, \mathcal{X}_i, \eta_j) = \text{Poisson pmf of } (Y_{ji} \mid \mathcal{X}_i, \eta_j) \text{ with parameter } \lambda_{ji}$$

$$\lambda_{ji} = \exp(\mathcal{X}_i \beta_j + \eta_j) \tag{4}$$

and

$g(\eta_1, \eta_2)$ = bivariate normal pdf with mean vector 0 and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \tag{5}$$

A Mata Function for 2D Integration in the ME Context

-- We have written a Mata function that implements Gauss-Legendre quadrature for approximating non-closed form 2-dimensional integrals like (2) in the ME context.

-- This function is called "`bivquadleg`" and is implemented in the following way:

```
doubleintegralvec =
    bivquadleg(&integrand(),limits1,limits2,wtsandabs)
```

where

"`integrand`" specifies the name of a Mata function for the relevant integrand.

(should be coded so as to accommodate $n^{*} \times R^2$ matrix arguments – where typically, $n^{*} = n$ [the sample size] and R is the number of abscissae and weights to be used for the quadrature).

A Mata Function for 2D Integration in the ME Context (cont'd)

```
doubleintegralvec =
   bivquadleg(&integrand(),limits1,limits2,wtsandabs)
```

"`limits1`" is an n*×2 matrix of integration limits (possibly observation-specific) for the first argument– first and second columns contain the lower and upper limits of integration, respectively.

"`limits2`" is similarly defined.

"`wtsandabs`" is a R×2 matrix of weights and abscissae to be used for the quadrature

"`doubleintegralvec`" function output -- n*×1 vector of integral values.

A Mata Function for 2D Integration in the ME Context (cont'd)

Prior to invoking bivquadleg, the requisite Gauss-Legendre quadrature weights and abscissae must be obtained using the function "`GLQwtsandabs`" which is called in the following way

```
wtsandabs = GLQwtsandabs(bivquadpts)
```

where "`bivquadpts`" is the number of weights and abscissae to be used for the quadrature.

Mata Implementation in the Bivariate Count-Valued SUR Context

$$q(\breve{\theta},\ Z_i) = \ln\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{(Y_{1i}\,|\,X_i,\,\eta_1)}(Y_{1i}, X_i, \eta_1) \times f_{(Y_{2i}\,|\,X_i,\,\eta_2)}(Y_{2i}, X_i, \eta_2)\ g(\eta_1, \eta_2)\ d\eta_1 d\eta_2\right] \qquad (3)$$

-- Recall that the `bivquadleg` function has four arguments:

1) The first argument of the `bivquadleg` function is the integrand function. In the above example this is

$$f_{(Y_{1i}\,|\,\mathcal{X}_i,\,\eta_1)}(Y_{1i}, \mathcal{X}_i, \eta_1)\ f_{(Y_{2i}\,|\,X_i,\,\eta_2)}(Y_{2i}, \mathcal{X}_i, \eta_2)\ g(\eta_1, \eta_2)$$

Mata Implementation in the Bivariate Count-Valued SUR Context (cont’d)

The code for this is

```
real matrix BivPoissNormIntegrand(xxu1,xxu2){
external y1
external y2
external xb1
external xb2
external sigmasq1
external sigmasq2
external sigma12
lambda1=exp(xb1:+xxu1)
lambda2=exp(xb2:+xxu2)
poisspart=poissonp(lambda1,y1):*poissonp(lambda2,y2)
SIGMA=sigmasq1,sigma12 \
           sigma12,sigmasq2
xxu=colshape(xxu1,1),colshape(xxu2,1)
factor=rowsum((xxu*invsym(SIGMA)):*xxu)
bivnormpart=(1:/(2:*pi():*sqrt(det(SIGMA))))/*
*/:*exp(-.5:*factor)
matbivnormpart=colshape(bivnormpart,cols(xxu1))
integrandvals=poisspart:*matbivnormpart
return(integrandvals)
```

Mata Implementation in the Bivariate Count-Valued SUR Context (cont'd)

The code with comments is

```
/***************************************************
** Mata Function to compute the integrand for the
** bivariate Poiss-Norm objective function
** (log-likelihood) to be used by the Mata
** moptimize procedure.
***************************************************/
real matrix BivPoissNormIntegrand(xxu1,xxu2){

/***************************************************
** Set necessary externals.
***************************************************/
external y1
external y2
external xb1
external xb2
external sigmasq1
external sigmasq2
external sigma12

/***************************************************
** Construct Bivariate-Poisson-Normal PMF
```

```
** (likelihood).
*************************************************/
/*************************************************
** Indexes for the Poisson part of the likelihood.
** note that xb1 and xb2 are (nx1) and xxu1 and
** xxu2 are nx(J^2) where J is the number of
** quadrature points
*************************************************/
lambda1=exp(xb1:+xxu1)
lambda2=exp(xb2:+xxu2)

/*************************************************
** Construct Poisson part of the likelihood.
** note that both of the factors are nx(J^2)
*************************************************/
poisspart=poissonp(lambda1,y1):*poissonp(lambda2,y2)

/*************************************************
** Construct Normal part of the likelihood.
*************************************************/
/*************************************************
** Construct SIGMA the 2x2 covariance matrix for
** the normal part of the likelihood.
*************************************************/
SIGMA=sigmasq1,sigma12 \
           sigma12,sigmasq2
```

```
/*********************************************
** Reshape and concatenate xxu1 and xxu2 as
** n*(J^2)x1 column vectors.
*********************************************/
xxu=colshape(xxu1,1),colshape(xxu2,1)

/*********************************************
** Calculate xxu*invsym(SIGMA)*xxu' for each row
** of xxu while stacking the results as a
** [n*(J^2)]x1 column vector.
*********************************************/
factor=rowsum((xxu*invsym(SIGMA)):*xxu)

/*********************************************
** Construct bivariate normal part of the
** likelihood which is at this point a [n*(J^2)]x1
** column vector.
*********************************************/
bivnormpart=(1:/(2:*pi():*sqrt(det(SIGMA))))/*
*/:*exp(-.5:*factor)

/*********************************************
** Reshape bivnormpart as an nx(J^2) matrix so as
** to conform to poisspart.
*********************************************/
```

```
matbivnormpart=colshape(bivnormpart,cols(xxu1))

/************************************************
** Construct the nx(J^2) matrix of Poisson-Norm
** likelihoods.
** Note that nx(J^2) is the order of the matrix
** that the "bivquadleg" bivariate quadrature
** function is designed to accept.
************************************************/
integrandvals=poisspart:*matbivnormpart
/************************************************
** Return result.
************************************************/
return(integrandvals)
```

Mata Implementation in the Bivariate Count-Valued SUR Context (cont'd)

-- Recall that the `bivquadleg` function has four arguments:

2) and 3) The second and third arguments of the `bivquadleg` function are the n×2 matrices of integration limits for the arguments of the integrand – first and second columns contain the lower and upper limits of integration – respectively.

Mata Implementation in the Bivariate Count-Valued SUR Context (cont'd)

In the above example, the code for this is

```
/*************************************************
** Integration limits for Biv-Poiss-Norm PMF.
*************************************************/
LIMITS=-8,8

/*************************************************
** Construct the obs x 2 matrix of
** observation-specific integration limits.
*************************************************/
limits=LIMITS#J(rows(X),1,1)
.
.
.
/*************************************************
** Construct the two n x 2 matrices of
** observation-specific integration limits.
*************************************************/
limits1=limits
limits2=limits
```

Mata Implementation in the Bivariate Count-Valued SUR Context (cont'd)

-- Recall that the `bivquadleg` function has four arguments:

4) The R×2 matrix of Gauss-Legendre weights and abscissae

The code for this is

```
/**************************************************
** Set the number of quadrature points for
** Bivariate-Poisson-Normal SUR regressions.
**************************************************/
bivquadpts=30

/**************************************************
** Compute the matrix of quadrature wts and
** abcissae.
**************************************************/
wtsandabs=GLQwtsandabs(bivquadpts)
```

Bivariate Count-Valued SUR: Simulation Sampling Design

$\mathcal{X}$ = [V   X] (two regressors), where V and X are both uniformly distributed with mean = 1 and variance = .25

β_j = [-1, -1, 0]' for j = 1, 2 (constant term is the last element)

$\sigma_j^2 = 1$

$\sigma_{12} = .9$

n = 50,000.

Bivariate Count-Valued SUR: Simulation Results

```
: moptimize_result_display(BivPoissNorm)

                                                Number of obs     =     50,000

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Y1           |
          x1 |  -.9773037    .021781   -44.87   0.000    -1.019994   -.9346137
          x2 |  -1.020545   .0218348   -46.74   0.000     -1.06334   -.9777494
       _cons |    .004543   .0295551     0.15   0.878    -.053384      .06247
-------------+----------------------------------------------------------------
Y2           |
          x1 |  -.9845058   .0220491   -44.65   0.000    -1.027721   -.9412903
          x2 |  -1.025328   .0220971   -46.40   0.000    -1.068637   -.9820183
       _cons |   -.017282   .0298589    -0.58   0.563    -.0758043    .0412403
-------------+----------------------------------------------------------------
sigmasq1     |
       _cons |   1.020365   .0271115    37.64   0.000     .9672275    1.073502
-------------+----------------------------------------------------------------
sigmasq2     |
       _cons |   1.036013   .0275469    37.61   0.000     .9820224    1.090004
-------------+----------------------------------------------------------------
sigma12      |
       _cons |   .8942762   .0209032    42.78   0.000     .8533066    .9352458
------------------------------------------------------------------------------
```

Some Notes

-- In general, nonlinear regression models of the relevant data generating process (DGP), of which (3) is an example, are typically intended for causal inference regarding the effect of a presumed cause (say X) on an outcome of interest (say Y). Terza (2019a) discusses the fact that results obtained from a such DGP-based regression models are causally interpretable only if model specification follows from a more primitive regression representation of the relationship between the X and the Y cast in a counterfactual (potential outcomes) framework.

Terza, J. V. (2019a). Regression-Based Causal Analysis from the Potential Outcomes Perspective. *Journal of Econometric Methods*, *0*(0). https://doi.org/10.1515/jem-2018-0030

Some Notes (cont'd)

1) We chose Gauss-Legendre (G-L) for our two-dimensional extension for the following reasons:

-- In comparison with other Gaussian quadratures, as detailed in Press et al. (2007) [viz. Chebyshev, Laguerre, Hermite, and Jacobi], Legendre appears to be the most generally applicable in that it does not require that the integrand conform to a particular class of functions.

-- Mata software for calculating G-L weights and abcissae is available. Thanks to Adrian Mander.

Press, W.W., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (2007): *Numerical Recipes, 3rd Edition*, Cambridge: Cambridge University Press.

Some Notes (cont'd)

-- I have already developed the uni-dimensional version of G-L quadrature for the M-Estimation context (Terza, 2019b)

2) The extension of G-L quadrature to 2D was motivated by our plan to develop a Mata implementation of the multivariate version of the count-valued SUR (CV-SUR) model.

-- The 2D quadrature will be sufficient for estimating the cross-equation covariances/correlations in a multivariate system, which can then be implemented in a nonlinear weighted least squares estimator of the system-wide regression parameters.

Terza, J. V. (2019b): " Mata Implementation of Gauss-Legendre Quadrature in the M-Estimation Context: Correcting for Sample-Selection Bias in a Generic Nonlinear Setting," 2019 Stata Conference 31, Stata Users Group.

Some Notes (cont'd)

-- Our CV-SUR model will be "plug-and-play" both with regard to its count-data components -- $f_{(Y_{ji} \mid X_i, \eta_j)}(Y_{ji}, X_i, \eta_j)$ for $j = 1, ..., J$ and its mixture component -- $g(\eta_1, ..., \eta_J)$.

-- The plug-and-play aspect of the former allows the implementation of dispersion-flexible specifications (Conway-Maxwell Poisson, Hyper Poisson, Restricted Generalized Poisson and old stand-bys like the Negative Binomial).

-- The plug-and-play aspect of the latter is facilitated by our implementation of G-L quadrature.

Application to Real Data (cont'd)

-- Medical-care demand by the elderly (age 66 or over) from the 1987 National Medical Expenditure Survey retrieved from Journal of Applied Econometrics Data Archive (Volume 12, Number 3, 1997).

-- The two outcomes are:

the number of visits to a physician in an office setting (OFP)

the number of visits to a nonphysician in an office setting (OFNP)

-- Regressor of primary interest: insurance coverage (our binary policy variable)

-- 16 control variables reflecting, health status, and other socioeconomic and demographic characteristics.

-- n = 4,406

Application to Real Data (cont'd)

These data weres analyzed by Deb & Trivedi (1997), Chib & Winkelmann (2001) and many others in the multivariate count literature.

Number of quadrature points = 30.

Partha Deb and Pravin K. Trivedi (1997): "Demand for Medical Care by the Elderly: A Finite Mixture Approach," *Journal of Applied Econometrics*, 12, 313-336.
Chib, S., & Winkelmann, R. (2001): "Markov Chain Monte Carlo Analysis of Correlated Count Data," *Journal of Business & Economic Statistics*, 19), 428–435. https://doi.org/10.1198/07350010152596673

Application to Real Data (cont'd)

Coefficients/Outcomes	# of Physician Office visits		# of Non-Physician Office visits	
	(Eq 1) Single Equation Poiss-Norm	(Eq 1) SUR Poiss-Norm using bivquadleg	(Eq 2) Single Equation Poiss-Norm	(Eq 2) SUR Poiss-Norm using bivquadleg
X (private insurance)	0.411 (9.233)	0.424 (9.529)	0.908 (8.636)	0.888 (8.981)
exclhlth	-0.355 (6.661)	-0.393 (6.672)	-0.297 (10.334)	-0.399 (12.991)
poorhlth	0.307 (8.757)	0.286 (8.981)	-0.384 (10.135)	-0.372 (12.818)
# of chronic diseases	0.242 (34.024)	0.226 (35.28)	0.130 (37.832)	0.173 (41.789)
adl difficulty	0.027 (9.703)	0.041 (10.482)	0.157 (14.269)	0.270 (17.442)
Midwest	0.098 (9.789)	0.071 (9.672)	0.572 (8.484)	0.253 (13.928)
Western US	-0.013 (10.168)	-0.001 (10.723)	0.737 (8.386)	0.520 (13.964)
age	0.143 (9.347)	0.136 (9.575)	1.175 (7.820)	0.916 (13.381)
race	-0.006 (14.877)	0.007 (15.598)	-0.052 (18.418)	-0.060 (20.951)
male	-0.147 (7.517)	-0.119 (8.189)	-0.239 (7.864)	-0.205 (8.839)

married	**-0.142 (12.339)**	**-0.151 (11.914)**	**-0.678 (15.325)**	**-0.611 (17.710)**
Years of schooling	**0.004 (11.338)**	**0.010 (11.584)**	**0.224 (15.3180**	**0.129 (17.911)**
Northeast	**0.028 (91.933)**	**0.022 (87.991)**	**0.042 (114.784)**	**0.037 (129.874)**
family income	**-0.003 (68.406)**	**0.003 (71.126)**	**0.011 (113.060)**	**0.020 (118.028)**
employed	**0.034 (8.497)**	**0.016 (9.116)**	**0.229 (9.698)**	**0.182 (12.320)**
medicaid	**0.387 (6.692)**	**0.344 (6.870)**	**-0.059 (7.375)**	**-0.341 (7.352)**
constant	**0.339 (1.893)**	**0.292 (1.941)**	**-3.096 (2.325)**	**-2.678 (2.538)**
sigma	**0.956 (55.151)**	**0.900 (33.93)**	**1.453 (56.08)**	**1.874 (39.74)**
rho	**N/A**	**0.312 (14.436)**	**N/A**	**0.312 (14.436)**