Driver PMAIN for DGL[FG][B]

David M. Gay

DGLG (and it's finite-difference variant DGLF) address robust, maximum-quasi-likelihood, and maximum-likelihood regression problems that involve nuisance parameters: for $x = \begin{bmatrix} \beta \\ \gamma \end{bmatrix}$,

$$f(x) := f(\beta,\gamma) = \sum_{i=1}^{n} \rho_i(\eta_i(\beta),\gamma),$$

where $\beta \in \mathbb{R}^p$ is the vector of parameters being estimated, $\eta_i : \mathbb{R}^p \to \mathbb{R}$ is a model for the ith observation (typically η_i has the form $\eta_i(\beta) = \eta(\beta, y_i)$, where y_i is the ith observation), and $\gamma \in \mathbb{R}^\nu$ is a vector of nuisance parameters. Use of DGLG and DGLF and of their simply-bounded counterparts, DGLGB and DGLFB, is summarized in the accompanying PORT-style description. This note gives details of my test program, `pmain.f`, and two subroutines, `rhpoil.f` (for RHO) and `brj.f` (for CALCRJ), that this test program uses. The latter include several choices of the r and ρ, described in detail below. The program is menu driven; menu item 1 causes the menu to be displayed. Currently the menu is:

```
1 = LIST MENU
2 = READ IV
3 = READ V
4 = READ ALG: 1 = GLG, 2 = GLF, 3 = GLGB, 4 = GLFB
5 = READ ALL OF X0
6 = COPY X TO X0
7 = START
8 = CONTINUE
9 = READ COMMANDS FROM SPECIFIED FILE
10 = READ PROBLEM
11 = READ RHO
12 = READ MODEL
13 = CHECK RHO DERIVATIVES
14 = READ P
15 = READ X0 COMPONENTWISE
16 = read new Y
17 = negate RHO (negative ==> use weights; see KW = 19)
18 = read KDIAG: 1 = from X*, 2 = from X0, 3 = both
19 = read KW: 1 = RHO", 2 = IRLS, 3 = score, 4 = deviance
20 = READ B (format i, b(1,i), b(2,i))
21,22 = Read,Show RHOI (componentwise)
23,24 = Read,Show RHOR        "
25 = Show range of RHOR components
26,27 = Show IV, V components
28 = Read and echo comment
```

`READ IV` reads pairs of integers (i, j) and sets IV(i) := j until $i \le 0$. This is for overriding default IV values.

`READ V` reads pairs of numbers (i, σ) (i an integer, σ a floating-point number) and sets V(i) := σ until $i \le 0$. This is for overriding default V values.

`READ ALG` reads an integer that tells which solver to invoke.

`READ ALL OF X0` lets you specify an alternate starting guess $x^\circ = \begin{bmatrix}\beta^\circ \\ \gamma^\circ\end{bmatrix}$.

`COPY X TO X0` copies the current x vector to x° for use in a subsequent `START`.

`START` starts the algorithm from scratch (from x°).

`CONTINUE` continues the iteration where it left off (from the current x). Presumably you have altered one or more IV or V components.

`READ COMMANDS FROM SPECIFIED FILE` must be followed by the name of an input file on the next input line; further commands are read from this new input file until it is exhausted, after which command reading resumes with the previous input file. These input diversions can be nested to a maximum depth of 5.

`READ PROBLEM` expects to read the name of a problem input file on the next input line. If this name is `-` (i.e., a single minus sign), then the problem is read from the current input file. The problem file, in turn, starts with

1. a one-line description of the data set;
2. a line containing 8 integers (format 8I4):
   - 2.1 n, the number of observations;
   - 2.2 p, the number of parameters, i.e., x components, including the nuisance parameters;
   - 2.3 k_η, the choice of η, described below;
   - 2.4 m, the number of extra data per observation (beyond y_i and c_i) to read;
   - 2.5 k_ρ the choice of ρ, described below;
   - 2.6 *dummy* (ignored; for compatibility with Frome's data format);
   - 2.7 *dummy* (ignored; for compatibility with Frome's data format);
   - 2.8 p_S (the number of non-nuisance parameters; if p_S is read as 0, then p_S is set to p).
3. For $k_\eta > 2$, the initial guess x°; for $k_\eta \le 2$ and $p_S < p$, the initial guess $x^\circ_{p_s+1, \ldots, p}$ for the nuisance parameters.
4. A `FORMAT` for reading the next n lines.

$*$. n data lines giving y_i, c_i, and $A_{i,j}$, $1 \le j \le m$. The matrix A^T thus defined is passed to BRJ (the CALCRJ subroutine) as UR, while the $(y_i,\ c_i)$ pairs are passed to RHPOIL (the RHO subroutine) as RHOR.

`READ RHO` causes a new k_ρ to be read from the next input line.

`READ MODEL` causes a new k_η to be read from the next input line.

`CHECK RHO DERIVATIVES` is a debugging item. I forget whether it copes with nuisance parameters.

`READ P` causes a new p to be read from the next input line.

`READ X0 COMPONENTWISE` reads pairs (i, σ) and sets $x^\circ{}_i := \sigma$ until $i \le 0$.

`read new Y` causes new $(y_i,\ c_i)$ values to be read (under the previously read `FORMAT`) from the next n input lines.

`negate RHO` causes $k_\rho := -k_\rho$; for some values of $|k_\rho|$(those for which I knew the appropriate w), $k_\rho < 0$ causes Roy's choice of the Gauss-Newton model, i.e., $J^T \mathrm{diag}(w) J$, to be used instead of $J^T \mathrm{diag}\left(\frac{\partial^2 \rho}{\partial \eta_i^2}\right) J$.

`read KDIAG` is for comparing one-step and fully iterated leave-one-out diagnostics with approximations based on a quadratic model (with a finite-difference Hessian approximation). [The real testing routine, louchk.c, is written in C and is not included in the Fortran-only version of this test program.]

`read KW` specifies the choice of diagonal matrix $\langle w \rangle$ used in the Gauss-Newton Hessian approximation $J^T \langle w \rangle J$.

`READ B` permits setting selected components of the array of simple bounds passed to DGLFB and DGLGB. Initially all lower bounds are $-\infty$ and all upper bounds are $+\infty$.

`Read,Show RHOI/RHOR (componentwise)` and `Show range of RHOR components` are occasionally useful for debugging or playing with block or fractional regression diagnostics.

`Show IV, V components` are mainly for debugging.

`Read and echo comment` may be useful in input scripts that solve several problems.

Models $\eta(\beta)$

Choices for k_η include:

(1) *Linear*: $\eta_i(\beta) = A_i\beta = \sum_{j=1}^{p_s} A_{i,j}\beta_j$.

(2) *Log-linear*: $\eta_i(\beta) = \exp(A_i\beta) = \exp(\sum_{j=1}^{p_s} A_{i,j}\beta_j)$.

(3) *Log of (4).*

(4) *Nonlinear Poisson model from Frome's PREG manual.*

(5) *Caesium dose effect model (2.6 of [Frome, 1984]).*

(6) *Lung cancer model (2.8 of [Frome, 1984]).*

(7) *Logistic of linear*: $\eta_i(\beta) = \dfrac{\exp(A_i\beta)}{1 + \exp(A_i\beta)}$.

(8) *Exp-linear*: $\eta_i(x) = \log(A_i\beta)$.

(9) *Model from p. 204–5 of [McCullagh and Nelder]*:

$$\eta_i(\beta) = \beta_1 + \beta_2 \log(A_{i,1}) + \frac{\beta_3 A_{i,2}}{\beta_4 + A_{i,2}}.$$

In this case, $m = 2$.

(10) *Model from p. 205–6 of [McCullagh and Nelder]*:

$$\eta_i(\beta) = \beta_1 + \beta_2 \log(A_{i,1} - \beta_5) + \frac{\beta_3 A_{i,2}}{\beta_4 + A_{i,2}},$$

where once again $m = 2$.

(11) *Model from p. 202 of [McCullagh and Nelder]*:

$$\eta_i(\beta) = \beta_1 + \sum_{j=1}^{3} \frac{\beta_{2j}}{\alpha_{i,j} + \beta_{2j+1}}.$$

Loss Functions $\rho_i(\eta_i, \gamma)$

Choices for k_ρ include:

(1) *IRLS Poisson*: $\mu = \eta$: $\rho_i(\eta) = c_i\eta - y_i\log(\eta)$, but we lie about $\dfrac{\partial^2\rho_i}{\partial\eta^2}$: we use the IRLS weights $w_i = c_i/\eta$ instead of $\dfrac{\partial^2\rho_i}{\partial\eta^2}$ for both RD$(i,1)$ and W(i).

(2) *Poisson*: Again $\rho_i(\eta) = c_i\eta - y_i\log(\eta)$, but now we supply the true $\dfrac{\partial^2\rho_i}{\partial\eta^2}$ for both RD$(i,1)$ and W(i).

(–2) *Poisson with IRLS GN model*: This is a mixture of (1) and (2): W(i) = w_i and RD$(i,1)$ = $\dfrac{\partial^2\rho_i}{\partial\eta^2}$.

(3) *Log Linear Poisson*: $\rho_i(\eta) = c_i \exp(\eta) - y_i \eta$.

(–3) *IRLS GN = No change*: $\mu = \exp(\eta)$ ==> $w_i = c_i \exp(\eta)$ ==> IRLS GN = ρ'' GN.

(4) *Square-root Linear for Poisson*: $\rho_i(\eta) = c_i \eta^2 - 2 y_i \log \eta$.

(–4) *IRLS*: $\mu = \eta^2$ ==> $w_i = 4 c_i$.

(5) *Binomial IRLS*: $\mu = \eta$: $\rho_i(\eta) = -y_i \log(\eta) - (c_i - y_i) \log(1 - \eta)$, with RD($i$,1) = W($i$) = IRLS weight = $\eta^{-1}(1 - \eta)^{-1}$.

(6) *Binomial*: ρ_i as for (5), but RD(i,1) = W(i) = $\dfrac{\partial^2 \rho_i}{\partial \eta^2}$.

(–6) *Binomial with IRLS GN model*: This is a mixture of (5) and (6), analogous to (–2).

(7) *Binomial Logistic*: $\rho_i(\eta) = c_i \log(1 + e^\eta) - y_i \eta$.

(–7) *IRLS GN = No change*: $\mu = \dfrac{\exp(\eta)}{1 + \exp(\eta)}$ ==> $w_i = \mu(1 - \mu) = \dfrac{e^\eta}{(1 + e^\eta)^2}$.

(8) *Probit*: (binomial with $\mu = \Phi(\eta)$, Φ = cumulative normal distribution):

$$\rho_i(\eta) = -y_i \log[\Phi(\eta)] - (c_i - y_i) \log[1 - \Phi(\eta)].$$

(–8) *IRLS Probit*: $\mu = \Phi(\eta)$ ==> $w_i = \dfrac{e^{-\eta^2}}{2\pi\mu(1 - \mu)}$.

(9) *Weibull*:

$$\rho_i(\eta) = -y_i \log[1 - \exp(-e^\eta)] + (c_i - y_i) e^\eta.$$

(–9) *IRLS Weibull*: $\mu = 1 - e^{-e^\eta}$ ==> $w_i = \dfrac{e^{2\eta}}{e^{e^\eta} - 1}$.

(10) *Gamma*: $\rho_i(\eta) = c_i (y_i \eta - \log(\eta))$.

(–10) *IRLS GN = No change*: $\mu = 1 / \eta$ ==> $w_i = \mu^2 = \eta^{-2}$ ==> IRLS GN = ρ'' GN.

(11) *Nelder and Pregibon Quasi-Likelihood, replicate weights*:

$$\rho_i(\eta,\gamma) = c_i \left[\tfrac{1}{2} \log[2\pi\phi y_i^\theta] + \left[\frac{y_i^{1-\theta} - \eta^{\psi(1-\theta)}}{\phi(1 - \theta)} \right] y_i + \frac{\eta^{\psi(2-\theta)} - y_i^{2-\theta}}{\phi(2 - \theta)} \right],$$

where $\phi = \gamma_1$, $\theta = \gamma_2$, and $\psi = \gamma_3$. (L'Hôpital's rule gives special forms for ρ_i when θ = 1 or 2.)

(–11) *IRLS*: $\mu = \eta^\psi$ ==> $w_i = c_i \psi^2 \eta^{(2 - \theta)\psi - 2}$.

(12) *Least Squares*: $\rho_i(\eta) = \tfrac{1}{2}(\eta - y_i)^2$.

(13) *Nelder and Pregibon Quasi-Likelihood, variance weights*:

$$\rho_i(\eta,\gamma) = \tfrac{1}{2} \log[2\pi\phi y_i^\theta / c_i] + c_i \left[\left[\frac{y_i^{1-\theta} - \eta^{\psi(1-\theta)}}{\phi(1 - \theta)} \right] y_i + \frac{\eta^{\psi(2-\theta)} - y_i^{2-\theta}}{\phi(2 - \theta)} \right],$$

where ϕ, θ, and ψ are as for (11).

(–13) : see (–11).