Optimization MT 1.0

for GAUSS™

Part Number: `007200`
Version 1.0
Documentation Revision: 1700 March 18, 2010

Contents

1 Installation

1.1 UNIX/Linux/Mac

If you are unfamiliar with UNIX/Linux/Mac, see your system administrator or system documentation for information on the system commands referred to below.

1.1.1 Download

1. Copy the `.tar.gz` or `.zip` file to `/tmp`.
2. If the file has a `.tar.gz` extension, unzip it using `gunzip`. Otherwise skip to step 3.

   `gunzip app_`*appname*`_`*vernum*`.`*revnum*`_UNIX.tar.gz`
3. `cd` to your **GAUSS** or **GAUSS Engine** installation directory. We are assuming `/usr/local/gauss` in this case.

   ```
   cd /usr/local/gauss
   ```

4. Use `tar` or `unzip`, depending on the file name extension, to extract the file.

   ```
   tar xvf /tmp/app_appname_vernum.revnum_UNIX.tar
   ```

   – or –

   ```
   unzip /tmp/app_appname_vernum.revnum_UNIX.zip
   ```

1.1.2 CD

1. Insert the Apps CD into your machine's CD-ROM drive.
2. Open a terminal window.
3. `cd` to your current **GAUSS** or **GAUSS Engine** installation directory. We are assuming `/usr/local/gauss` in this case.

   ```
   cd /usr/local/gauss
   ```

4. Use `tar` or `unzip`, depending on the file name extensions, to extract the files found on the CD. For example:

   ```
   tar xvf /cdrom/apps/app_appname_vernum.revnum_UNIX.tar
   ```

   – or –

   ```
   unzip /cdrom/apps/app_appname_vernum.revnum_UNIX.zip
   ```

   However, note that the paths may be different on your machine.

1.2 Windows

1.2.1 Download

Unzip the `.zip` file into your **GAUSS** or **GAUSS Engine** installation directory.

1.2.2 CD

1. Insert the Apps CD into your machine's CD-ROM drive.

2. Unzip the `.zip` files found on the CD to your **GAUSS** or **GAUSS Engine** installation directory.

1.2.3 64-Bit Windows

If you have both the 64-bit version of **GAUSS** and the 32-bit Companion Edition installed on your machine, you will need to install the application in both **GAUSS** installation directories.

1.3 Difference Between the UNIX and Windows Versions

- If the functions can be controlled during execution by entering keystrokes from the keyboard, it may be necessary to press ENTER after the keystroke in the UNIX version.

2 Getting Started

GAUSS 10 and the **GAUSS** Run-Time library 10+ are required to use these routines. See **_rtl_ver** in src/gauss.dec.

The **Optimization MT** version number is stored in a global variable:

_optmt_ver 3×1 matrix, the first element contains the major version number, the second element the minor version number, and the third element the revision number.

If you call for technical support, you may be asked for the version of your copy of **Optimization MT**.

2.0.1 README Files

The file **README.optmt** contains any last minute information on the **Optimization MT** procedures. Please read it before using them.

3 Optimization MT

written by

Ronald Schoenberg

This module contains a set of procedures for the solution of the optimization problem

3.1 Getting Started

GAUSS 10+ is required to use these routines.

3.1.1 Setup

In order to use the procedures in the *Optimization MT* or **Optmt** Module, the **Optmt** library must be active. This is done by including `optmt` in the **LIBRARY** statement at the top of your program or command file:

```
library optmt,pgraph;
```

This enables **GAUSS** to find the **Optmt** procedures. The statement

```
#include optmt.sdf
```

is also required. It sets the definitions of the structures used by **Optmt**

The version number of each module is stored in a global variable:

_optmt_ver 3×1 matrix, the first element contains the major version number of the **Optmt** Module, the second element the minor version number, and the third element the revision number.

If you call for technical support, you may be asked for the version number of your copy of this module.

3.2 Special Features in Optimization MT

3.2.1 Structures

In **Optmt** the same procedure computing the objective function will be used to compute analytical derivatives as well if they are being provided. Its return argument is a

optmtResults structure with three members, a scalar objective function evaluation, an optional Kx1 vector of first derivatives, and an optional KxK matrix of second derivatives. Of course the derivatives are optional, or even partially optional, i.e., you can compute a subset of the derivatives if you like and the remaining will be computed numerically. This procedure will have an additional argument which tells the function which to compute, the log-likelihood or objective, the first derivatives, or the second derivatives, or all three. This means that calculations in common won't have to be redone.

Optmt uses the **DS** and **PV** structures that are now in use in the **GAUSS** Run-Time Library. The **DS** structure is completely flexible, allowing you to pass anything you can think of into your procedure. The **PV** structure revolutionizes how you pass the parameters into the procedure. No more do you have to struggle to get the parameter vector into matrices for calculating the function and its derivatives, trying to remember, or figure out, which parameter is where in the vector. If your objective functions uses matrices or arrays,you can store them directly into the **PV** structure, and remove them as matrices or arrays with the parameters already plugged into them. The PV structure can handle matrices and arrays where some of their elements are fixed and some free. It remembers the fixed parameters and knows where to plug in the current values of the free parameters. It can handle symmetric matrices where parameters below the diagonal are repeated above the diagonal.

There will no longer be any need to use global variables. Anything the procedure needs can be passed into it through the **DS** structure. And these new applications will use control structures rather than global variables. This means, in addition to thread safety, that it will be straightforward to nest calls to **Optmt** inside of a call to **Optmt**, not to mention Run-Time Library functions like **QNewtonmt**, **QProgmt**, and **EQsolvemt**.

3.2.2 Simple Bounds

Bounds may be placed on parameters. This can be very important for models with a limited parameter space outside of which the log-likelihood is not defined.

3.2.3 Threading

if you have a multi-core processor in your computer, you may take advantage of this capability by selecting threading. This is done by setting the `useThreads` member of the **cmlmtControl** instance:

```
struct optmtControl c0;
c0 = optmtControlCreate;
c0.useThreads = 1;
```

An important advantage of threading occurs in computing numerical derivatives. if the derivatives are computed numerically threading will significantly decrease the time of computation.

3.3 The Objective Function

The procedure **Optmt** finds values for the parameters in θ such that $F(\theta)$ is minimized.

Optmt has been designed to make the specification of the function and the handling of the data convenient. The user supplies a procedure that computes F given the parameters in θ,

3.4 Algorithm

Optmt includes three descent methods, BFGS, DFP, and Newton. In these methods the parameters are updated in a series of iterations beginning with a starting values that you provide. Let θ_t be the current parameter values. Then the succeeding values are

$$\theta_{t+1} = \theta_t + \rho\delta$$

where δ is a $K \times 1$ *direction* vector, and ρ a scalar *step length*.

Direction

Define

$$\begin{aligned}\Sigma(\theta) &= \frac{\partial^2 F}{\partial\theta\partial\theta'}\\ \Psi(\theta) &= \frac{\partial F}{\partial\theta}\end{aligned}$$

The direction, δ is the solution to

$$\Sigma(\theta_t)\delta = \Psi(\theta_t)$$

This solution requires that Σ be positive semi-definite.

Line Search

Define the merit function

$$m(\theta) = L - \sum_{\ell} \lambda_\ell h_\ell(\theta)$$

h_ℓ is the ℓ-th bounds constraint and λ_ℓ the Lagrangean coefficient of the ℓ-th bounds constraint.

The line search finds a value of ρ that minimizes or decreases $m(\theta_t + \rho\delta)$.

3.4.1 The Secant Algorithms

The Hessian may be very expensive to compute at every iteration, and poor start values may produce an ill-conditioned Hessian. For these reasons alternative algorithms are provided in **Optmt** for updating the Hessian rather than computing it directly at each iteration. These algorithms, as well as step length methods, may be modified during the execution of **Optmt**.

Beginning with an initial estimate of the Hessian, or a conformable identity matrix, an update is calculated. The update at each iteration adds more "information" to the estimate of the Hessian, improving its ability to project the direction of the descent. Thus after several iterations the secant algorithm should do nearly as well as Newton iteration with much less computation.

There are two basic types of secant methods, the BFGS (Broyden, Fletcher, Goldfarb, and Shanno), and the DFP (Davidon, Fletcher, and Powell). They are both rank two updates, that is, they are analogous to adding two rows of new data to a previously computed moment matrix. The Cholesky factorization of the estimate of the Hessian is updated using the functions **CHOLUP** and **CHOLDN**.

Secant Methods (BFGS and DFP)

BFGS is the method of Broyden, Fletcher, Goldfarb, and Shanno, and DFP is the method of Davidon, Fletcher, and Powell. These methods are complementary (Luenberger 1984, page 268). BFGS and DFP are like the NEWTON method in that they use both first and second derivative information. However, in DFP and BFGS the Hessian is approximated, reducing considerably the computational requirements. Because they do not explicitly calculate the second derivatives they are sometimes called *quasi-Newton* methods. While it takes more iterations than the NEWTON method, the use of an approximation produces a gain because it can be expected to converge in less overall time (unless analytical second derivatives are available in which case it might be a toss-up).

The secant methods are commonly implemented as updates of the *inverse* of the Hessian.

This is not the best method numerically for the BFGS algorithm (Gill and Murray, 1972). This version of **Optmt**, following Gill and Murray (1972), updates the Cholesky factorization of the Hessian instead, using the functions **CHOLUP** and **CHOLDN** for BFGS. The new direction is then computed using **CHOLSOL**, a Cholesky solve, as applied to the updated Cholesky factorization of the Hessian and the gradient.

3.4.2 Line Search Methods

Given a direction vector d, the updated estimate of the parameters is computed

$$\theta_{t+1} = \theta_t + \rho\delta$$

where ρ is a constant, usually called the *step length*, that increases the descent of the function given the direction. **Optmt** includes a variety of methods for computing ρ. The value of the function to be minimized as a function of ρ is

$$m(\theta_t + \rho\delta)$$

Given θ and d, this is a function of a single variable ρ. Line search methods attempt to find a value for ρ that decreases m. STEPBT is a polynomial fitting method, BRENT and HALF are iterative search methods. A fourth method called ONE forces a step length of 1. The default line search method is STEPBT. If this, or any selected method, fails, then BRENT is tried. If BRENT fails, then HALF is tried. If all of the line search methods fail, then a random search is tried, provided the `RandRadius` member of the **optmtControl** instance is greater than zero which it is by default.

STEPBT

STEPBT is an implementation of a similarly named algorithm described in Dennis and Schnabel (1983). It first attempts to fit a quadratic function to $m(\theta_t + \rho\delta)$ and computes a ρ

that minimizes the quadratic. If that fails it attempts to fit a cubic function. The cubic function more accurately portrays the F which is not likely to be very quadratic, but is, however, more costly to compute. STEPBT is the default line search method because it generally produces the best results for the least cost in computational resources.

BRENT

This method is a variation on the *golden section* method due to Brent (1972). In this method, the function is evaluated at a sequence of test values for ρ. These test values are determined by extrapolation and interpolation using the constant, $(\sqrt{5} - 1)/2 = .6180....$ This constant is the inverse of the so-called "golden ratio" $((\sqrt{5} + 1)/2 = 1.6180...$ and is why the method is called a golden section method. This method is generally more efficient than STEPBT but requires significantly more function evaluations.

HALF

This method first computes $m(x + d)$, i.e., sets $\rho = 1$. If $m(x + d) < m(x)$ then the step length is set to 1. If not, then it tries $m(x + .5d)$. The attempted step length is divided by one half each time the function fails to decrease, and exits with the current value when it does decrease. This method usually requires the fewest function evaluations (it often only requires one), but it is the least efficient in that it is not very likely to find the step length that decreases m the most.

3.5 Bounds

To specify bounds, the lower and upper bounds respectively are entered in the first and second columns of a matrix that has the same number of rows as the parameter vector. This matrix is assigned to the `Bounds` member of an instance of a **optmtControl** structure.

If the bounds are the same for all of the parameters, only the first row is necessary.

To bound four parameters:

```
struct optmtControl ctl;
ctl = optmtControlCreate;
ctl.Bounds = { -10 10,
               -10  0,
                 1 10,
                 0  1 };
```

Suppose all of the parameters are to be bounded between -50 and +50, then,

```
ctl.Bounds = { -50 50 };
```

is all that is necessary.

3.6 The Optmt Procedure

The call to **Optmt** has four input arguments and one output argument.

3.6.1 First Input Argument: Pointer to Procedure

The first input argument is the pointer to the procedure computing the log-likelihood function and optionally the gradient and/or Hessian. See Section 3.7 for details.

3.6.2 Second Input Argument: PV parameter Instance

The **GAUSS** Run-Time Library contains special functions that work with the **PV** structure. They are prefixed by "pv" and defined in `pv.src`. These functions store matrices and

arrays with parameters in the structure, and retrieve the original matrices and arrays along with various kinds of information about the parameters and parameter vector from it.

The advantage of the **PV** structure is that it permits you to retrieve the parameters in the form of matrices and/or arrays ready for use in calculating your log-likelihood. The matrices and arrays are defined in your command file when the start values are set up. It isn't necessary that a matrix or array be completely free parameters to be estimated. There are **pvPack** functions that take mask arguments defining what is a parameter versus what is a fixed value. There are also functions for handling symmetric matrices where the parameters below the diagonal are duplicated above the diagonal.

For example, a PV structure is created in your command file:

```
struct PV p;
p = pvCreate;    // creates default structure

garch = { .1, .1, .1 };
p = pvPack(p,garch,"garch");
```

A matrix or array in the model may contain a mixture of fixed values along with parameters to be estimated. This type of matrix or array uses **pvPackm** which has an additional argument, called a "mask", strictly conformable to the input matrix or array indicating which elements are fixed (the corresponding element in the mask is zero) or being estimated (the corresponding element in the mask is nonzero). For example,

```
struct PV p;
p = pvCreate;

b = { 1.0  0.0  0.0,
      0.5  1.0  0.2,
      0.3  0.0  1.0 };

b_mask = { 0  0  0,
```

```
                1  0  1,
                1  0  1 };

p = pvPackm(p,b,"beta",b_mask);
```

In this case there are four free parameters to be estimated, b_{21}, b_{23}, b_{31}, and b_{33}. b_{11} and b_{22} are fixed to `1.0`, and b_{22}, b_{23}, and b_{32} are fixed to 0.0.

pvPacks "packs" a symmetric matrix into the **PV** structure in which only the lower left portion of the matrix contains independent parameters while the upper left is duplicated from the lower left. The following packed matrix contains three nonredundant parameters. When this matrix is unpacked it will contain the upper nonredundant portion of the matrix equal to the lower portion.

```
vc = { 1.2  0.4,
       0.4  2.1  };
p = pvPacks(p,vc,"phi");  // pack symmetric matrix
```

Suppose that you wish to specify a correlation matrix in which only the correlations are free parameters. You would then use **pvPacksm**.

```
cor = { 1.0  0.2,
        0.2  1.0 };
msk = { 0  1,
        1  0 };
pv = pvPacksm(p,cor,msk,"R");
```

Some computation speedup can be achieved by packing and unpacking by number rather than name. Each packing function has a version with an **i** suffix that packs by number. Then **pvUnpack** can be used with that number:

```
garch = { .1, .1, .1 };
p = pvPacki(p,garch,"garch",1);
```

which is unpacked using its number

```
g0 = pvUnpack(1);
```

3.6.3 Third Input Argument: DS Data Instance

The DS structure, or "data" structure, is a very simple structure. It contains a member for each **GAUSS** data type. This is its definition (see `ds.sdf` in the **GAUSS** `src` subdirectory):

```
struct DS {
            scalar type;
            matrix dataMatrix;
            array  dataArray;
            string dname;
            string array vnames;
          };
```

Data in Matrices or Arrays

If you are passing your data in as matrices or arrays, you can set the data structure in any way you want, except that the `dname` member of the first element of the data structure must be a null string. **Optmt** will pass this instance or a matrix of instances to your log-likelihood procedure untouched. For example

```
struct DS d0;
```

```
d0 = reshape(dsCreate,2,1);
d0[1].DataMatrix = y;
d0[2].DataMatrix = x;
```

3.6.4 Fourth Input Argument: optmtControl Instance

The members of the **optmtControl** instance determine everything about the optimization. For example, suppose you want **Optmt** to stop after 100 iterations:

```
struct optmtControl c0;
c0 = optmtControlCreate;

c0.maxIters = 100;
```

The **optmtControlCreate** procedure sets all of the defaults. The default values for all the members of a **optmtControl** instance can be found in that procedure, located at the top of `optmtutil.src` in the **GAUSS** `src` subdirectory.

3.7 The Objective Procedure

Optmt requires that you write a procedure computing the value of the objective function. The output from this procedure is a **modelResults** structure containing the value and optionally the first and second derivatives of the objective function with respect to the parameters. There are three input arguments to this procedure

1. instance of a **PV** structure containing parameter values
2. instance of a **DS** structure containing data
3. indicator vector

and one return argument

1. instance of a **modelResults** structure containing computational results.

3.7.1 First Input Argument: PV Parameter Instance

This argument contains the parameter matrices and arrays that you need for computing the objective function and (optionally) derivatives. The **pvUnpack** function retrieves them from the **PV** instance.

```
proc lpr(struct PV p, struct DS d, ind);
  local beta, gamma;
  beta = pvUnpack("beta");
  gamma = pvUnpack("gamma");
  .
  .
  .
endp;
```

You may have decided to speed the program up a bit by packing the matrices or arrays using the "i" pack functions, **pvPacki**, **pvPackmi**, **pvPacksi**, etc., You can then unpack the matrices and arrays with the integers used in packing them:

```
proc lpr(struct PV p, struct DS d, ind);
  local beta, gamma;
  beta = pvUnpack(1);
  gamma = pvUnpack(2);
  .
  .
  .
endp;
```

where it has been assumed that they've been packed accordingly:

```
struct PV p;
p = pvCreate;

p = pvPacki(p,1.|.1,"beta",1);
p = pvPacksi(p,(1~0)|(0~1),"gamma",2);
```

3.7.2 Second Input Argument: DS Data Instance

Optmt passes the **DS** instance you have constructed completely untouched. You can, therefore, design this instance completely for your convenience in computing the objective function and optionally its derivatives.

For example, you can write a general objective function procedure that computes a variety of objective functions, e.g., a probit and a logit. Then you can set the `Type` member of a **DS** instance to a value in your command file that chooses which to compute for that run.

In your command file

```
struct DS d;
d = dsCreate;
d.Type = 1;
d.dataMatrix = z;
```

and in your objective function procedure

```
proc lpr(struct PV p, struct DS d, ind);
    .
    .
    .
```

```
    if d.type == 1;  // compute probit objective function
      .
      .
    elseif d.type == 2;  // compute logit
      .
      .
    endif;
  .
  .
  .
endp;
```

3.7.3 Third Input Argument: Indicator Vector

The third argument is a vector with elements set to zero or one, indicating whether or not function, first derivatives, or second derivatives are to be computed.

1st element if nonzero, the function is to be computed.

2nd element if nonzero, the first derivatives are to be computed.

3rd element if nonzero, the second derivatives are to be computed.

The second and third elements associated with the first and second derivatives are optional.

For example,

```
proc logl( struct PV p0, struct DS d0, ind );
    local b0,b,y,x;
    b0 = pvUnpack(p0,"b0");
    b = pvUnpack(p0,"beta");
    y = d0[1].DataMatrix;
```

```
    x = d0[2].DataMatrix;

    struct modelResults mm;
    if ind[1];  // compute objective function
       mm.Function = ....
    endif;
    if ind[2];  // compute optional first derivatives
       mm.Gradient = ....
    endif;
    if ind[3];  // compute optional second derivatives
       mm.Hessian = ....
    endif;
    retp(mm);
endp;
```


If `mm.Gradient` and `mm.Hessian` are not set, they will be computed numerically by **Optmt**.

3.7.4 Output Argument: modelResults Instance

The return argument for your objective function procedure is an instance of a **modelResults** structure. The members of this structure are

1 scalar objective function

Function scalar objective function

Gradient 1 × K vector of first derivatives (optional) item[Hessian] K × K matrix of second derivatives (optional)

3.7.5 Examples

```
proc fct(struct PV p, struct DS d, ind);
    local x,h0;
    struct modelResults mm;

    x = pvUnpack(p,"x");

    if ind[1];
        mm.function = 100*(x[2]-x[1]^2)^2 + (1-x[1])^2 +
          90*(x[4]-x[3]^2)^2 + (1-x[3])^2 + 10.1*((x[2]-1)^2 +
          (x[4]-1)^2) + 19.8*(x[2]-1)*(x[4]-1);
    endif;
    if ind[2];
        local a,b;
        a = x[2] - x[1]^2;
        b = x[4] - x[3]^2;
        mm.gradient = zeros(4,1);
        mm.gradient[1] = -2*(200*x[1]*a + 1 - x[1]);
        mm.gradient[2] = 2*(100*a + 10.1*(x[2]-1) + 9.9*(x[4]-1));
        mm.gradient[3] = -2*(180*x[3]*b + 1 - x[3]);
        mm.gradient[4] = 2*(90*b + 10.1*(x[4]-1) + 9.9*(x[2]-1));
    endif;
    if ind[3];
        mm.hessian = d.dataMatrix;
        mm.hessian[1,1] = -2*(200*(x[2]-x[1]^2) - 400*x[1]^2 - 1);
        mm.hessian[1,2] = -400*x[1];
        mm.hessian[2,1] = mm.hessian[1,2];
        mm.hessian[3,3] = -2*(180*(x[4]-x[3]^2) - 360*x[3]^2 - 1);
        mm.hessian[4,3] = -360*x[3];
        mm.hessian[3,4] = mm.hessian[4,3];
     endif;
     retp(mm);
endp;
```

```
proc  qfct(struct PV p, struct DS d, ind);
    local x,q,b;
    struct modelResults mm;

    x = pvUnpack(p,"x");
    q = d[1].dataMatrix;
    b = d[2].dataMatrix;

    if ind[1];
        mm.function = .5*x'*q*x - x'b;
    endif;
    retp(mm);
endp;
```


3.8 Managing Optimization

The critical elements in optimization are scaling, starting point, and the condition of the model. When the data are scaled, the starting point is reasonably close to the solution, and the data and model go together well, the iterations converge quickly and without difficulty.

For best results therefore, you want to prepare the problem so that model is well-specified, the data scaled, and that a good starting point is available.

The tradeoff among algorithms and step length methods is between speed and demands on the starting point and condition of the model. The less demanding methods are generally time consuming and computationally intensive, whereas the quicker methods (either in terms of time or number of iterations to convergence) are more sensitive to conditioning and quality of starting point.

3.8.1 Scaling

For best performance, the diagonal elements of the Hessian matrix should be roughly equal. If some diagonal elements contain numbers that are very large and/or very small with respect to the others, **Optmt** has difficulty converging. How to scale the diagonal elements of the Hessian may not be obvious, but it may suffice to ensure that the constants (or "data") used in the model are about the same magnitude.

3.8.2 Condition

The specification of the model can be measured by the condition of the Hessian. The solution of the problem is found by searching for parameter values for which the gradient is zero. If, however, the Jacobian of the gradient (i.e., the Hessian) is very small for a particular parameter, then **Optmt** has difficulty determining the optimal values since a large region of the function appears virtually flat to **Optmt**. When the Hessian has very small elements, the inverse of the Hessian has very large elements and the search direction gets buried in the large numbers.

Poor condition can be caused by bad scaling. It can also be caused by a poor specification of the model or by bad data. Bad models and bad data are two sides of the same coin. If the problem is highly nonlinear, it is important that data be available to describe the features of the curve described by each of the parameters. For example, one of the parameters of the Weibull function describes the shape of the curve as it approaches the upper asymptote. If data are not available on that portion of the curve, then that parameter is poorly estimated. The gradient of the function with respect to that parameter is very flat, elements of the Hessian associated with that parameter is very small, and the inverse of the Hessian contains very large numbers. In this case it is necessary to respecify the model in a way that excludes that parameter.

3.8.3 Starting Point

When the model is not particularly well-defined, the starting point can be critical. When the optimization doesn't seem to be working, try different starting points. A closed form solution may exist for a simpler problem with the same parameters. For example, ordinary least squares estimates may be used for nonlinear least squares problems or nonlinear regressions like probit or logit. There are no general methods for computing start values and it may be necessary to attempt the estimation from a variety of starting points.

3.8.4 Example


```
library optmt;
#include optmt.sdf

struct DS d0;
d0 = reshape(d0,2,1);

d0[1].dataMatrix =
    { 0.78 -0.02 -0.12 -0.14,
     -0.02  0.86 -0.04  0.06,
     -0.12 -0.04  0.72 -0.08,
     -0.14  0.06 -0.08  0.74 };

d0[2].dataMatrix = { 0.76, 0.08, 1.12, 0.68 };

struct PV p0;
p0 = pvPack(pvCreate,1|1|1|1,"x");

struct optmtControl c0;
c0 = optmtControlCreate;


c0.bounds = { 0 1,
```

```
                0 1,
                1 2,
                1 2 };


proc  qfct(struct PV p, struct DS d, ind);
    local x,q,b;
    struct modelResults mm;

    x = pvUnpack(p,"x");
    q = d[1].dataMatrix;
    b = d[2].dataMatrix;

    if ind[1];
        mm.function = .5*x'*q*x - x'b;
    endif;
    retp(mm);

endp;

output file = optmt1.out reset;


struct optmtResults out;
out = optmt(&qfct,p0,d0,c0);

call optmtPrt(out);

output off;
```

and the output looks like this:

```
============================================================
 Optmt Version 1.0.0                  12/31/2009   3:14 pm
============================================================

return code =    0
normal convergence
  Parameters    Estimates   Gradient
------------------------------------------------------------
x[1,1]   1.0000     -0.3892
x[2,1]   0.1126      0.0000
x[3,1]   1.8731      0.0000
x[4,1]   1.3015      0.0000


Number of iterations    3
Minutes to convergence     0.00000

      0.00000000       0.38922812
      0.00000000       0.00000000
      0.00000000       0.00000000
      0.00000000       0.00000000
```

If the Lagrangeans are “empty” matrices, the associated constraints are not active. If they are zeros but not “empty” matrices, then they are still inactive at the solution but were active at some point during the iterations.

3.9 Run-Time Switches

If the user presses **H** during the iterations, a help table is printed to the screen which describes the run-time switches. By this method, important global variables may be modified during the iterations. The case may also be ignored, that is, either upper or lower case letters suffice.

A Change Algorithm
C Force Exit
G Toggle **GradMethod**
H Help Table
O Set **PrintIters**
S Set line search method
V Set **Tol**

Keyboard polling can be turned off completely by setting the `disableKey` member of the **optmtControl** instance to a nonzero value.

3.10 Optmt Structures

3.10.1 optmtControl

matrix	Bounds
matrix	Algorithm
matrix	Switch
matrix	LineSearch
matrix	Active
matrix	NumObs
matrix	MaxIters
matrix	Tol
matrix	Weights
matrix	CovParType
matrix	Alpha
matrix	FeasibleTest
matrix	MaxTries
matrix	RandRadius
matrix	GradMethod
matrix	HessMethod
matrix	GradStep
matrix	HessStep
matrix	GradCheck
matrix	State
string	Title
scalar	PrintIters
matrix	DisableKey
matrix	Select
matrix	Center
matrix	Increment
matrix	Width

matrix	NumCat
string	BootFileName
string	BayesFileName
matrix	BayesAlpha
scalar	PriorProc
matrix	NumSamples
matrix	MaxTime
matrix	MaxBootTime

3.10.2 optmtResults

struct	PV Par
scalar	Fct
struct	Lagrange
scalar	Retcode
string	ReturnDescription
matrix	CovPar
string	CovParDescription
matrix	NumObs
matrix	Hessian
matrix	Xproduct
matrix	Waldlimits
matrix	Inversewaldlimits
matrix	Bayeslimits
matrix	Profilelimits
matrix	Bootlimits
matrix	Gradient
matrix	NumIterations
matrix	ElapsedTime
matrix	Alpha
string	Title

3.10.3 modelResults

matrix	Function
matrix	Gradient
matrix	Hessian
array	Hessianw
matrix	NumObs

3.11 Error Handling

3.11.1 Return Codes

The `Retcode` member of an instance of a **optmtResults** structure, which is returned by **Optmt**, contains a scalar number that contains information about the status of the iterations upon exiting **Optmt**. The following table describes their meanings:

0	normal convergence
1	forced exit
2	maximum iterations exceeded
3	function calculation failed
4	gradient calculation failed
5	Hessian calculation failed
6	line search failed
7	function cannot be evaluated at initial parameter values
8	error with gradient
10	secant update failed
11	maximum time exceeded
12	error with weights
16	function evaluated as complex
20	Hessian failed to invert
34	data set could not be opened

3.11.2 Error Trapping

Setting the `PrintIters` member of an instance of a **optmtControl** structure to zero turns off all printing to the screen. Error codes, however, still are printed to the screen unless error trapping is also turned on. Setting the trap flag to 4 causes **Optmt** *not* to send the messages to the screen:

```
trap 4;
```

Whatever the setting of the trap flag, **Optmt** discontinues computations and returns with an error code. The trap flag in this case only affects whether messages are printed to the screen or not. This is an issue when the **Optmt** function is embedded in a larger program, and you want the larger program to handle the errors.

3.12 References

1. Amemiya, Takeshi, 1985. *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

2. Brent, R.P., 1972. *Algorithms for Minimization Without Derivatives*. Englewood Cliffs, NJ: Prentice-Hall.

3. Dennis, Jr., J.E., and Schnabel, R.B., 1983. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ: Prentice-Hall.

4. Fletcher, R., 1987. *Practical Methods of Optimization*. New York: Wiley.

5. Gill, P. E. and Murray, W. 1972. "Quasi-Newton methods for unconstrained optimization." *J. Inst. Math. Appl.*, 9, 91-108.

6. Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl and T.C. Lee. 1988. *Introduction to the Theory and Practice of Econometrics.* 2nd Edition. New York:Wiley.

7. Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl and T.C. Lee. 1985. *The Theory and Practice of Econometrics.* 2nd Edition. New York:Wiley.

Optmt Reference 4

Optmt

PURPOSE Computes estimates of parameters of a constrained maximum likelihood function.

LIBRARY **optmt**

FORMAT *out* = **Optmt(&***modelProc*,*par*,*data*,*ctl*);

INPUT *&modelProc* a pointer to a procedure that returns the value of the objective function

par instance of a **PV** structure containing start values for the parameters constructed using the **pvPack** functions.

data instance or matrix of instances of a **DS** structure containing data. it is passed to the user-provided procedure pointed at by `&fct` to be used in the objective function. There are two cases,

1 a scalar or vector of **DS** instances passed to **Optmt** are passed unchanged to the user-provided objective procedure. In this case the structure can be a scalar, vector, or matrix of **DS** instances, and all members of all the instances can be set at the discretion of the programmer, **except** that the **dname** member of the [1,1] element of the structure must be a null string.

ctl an instance of a **optmtControl** structure. Normally an instance is initialized by calling **optmtCreate** and members of this instance can be set to other values by the user. For an instance named *ctl*, the members are:

ctl`.Bounds` 1×2 or $K \times 2$ matrix, bounds on parameters. If 1×2 all parameters have same bounds. Default = { -1e256 1e256 }.

ctl`.Algorithm` scalar, descent algorithm.

1 BFGS (default)

2 DFP

3 Newton

ctl`.Switch` 4×1 or 4×2 vector, controls algorithm switching:
if 4×1:

ctl.Switch[1] algorithm number to switch to.

ctl.Switch[2] **Optmt** switches if function changes less than this amount.

ctl.Switch[3] **Optmt** switches if this number of iterations is

exceeded.

ctl.Switch[4] **Optmt** switches if line search step changes less than this amount.

else if 4×2 **Optmt** switches between the algorithm in column 1 and column 2. Default = { 1 3, .0001 .0001, 10 10, .0001 .0001 }.

ctl.LineSearch scalar, sets line search method.

- **1** STEPBT (quadratic and cubic curve fit) (default)
- **2** Brent's method
- **3** half
- **4** Wolfe's condition

ctl.Active $K \times 1$ vector, set K-th element to zero to fix it to start value. Use the **GAUSS** function **pvGetIndex** to determine where parameters in the **PV** structure are in the vector of parameters. Default = {.}, all parameters are active.

ctl.MaxIters scalar, maximum number of iterations. Default = 10000.

ctl.Tol scalar, convergence tolerance. Iterations cease when all elements of the direction vector are less than this value. Default = $1e-5$.

ctl.FeasibleTest scalar, if nonzero, parameters are tested for feasibility before computing function in line search. If function is defined outside inequality boundaries, then this test can be turned off. Default = 1.

ctl.MaxTries scalar, maximum number of attemps in random search. Default = 100.

ctl.RandRadius scalar, If zero, no random search is attempted. If nonzero, it is the radius of the random search. Default = .001.

ctl.GradMethod scalar, method for computing numerical gradient.

- **0** central difference
- **1** forward difference (default)
- **2** backward difference

ctl.HessMethod scalar, method for computing numerical Hessian.

- **0** central difference
- **1** forward difference (default)
- **2** backward difference

ctl.GradStep scalar or K × 1, increment size for computing numerical gradient. If scalar, stepsize will be value times parameter estimates for the numerical gradient. If K × 1, the step size for the gradient will be the elements of the vector, i.e., it will not be multiplied times the parameters.

ctl.HessStep scalar or K × 1, increment size for computing numerical Hessian. If scalar, stepsize will be value times parameter estimates for the numerical Hessian. If K × 1, the step size for the gradient will be the elements of the vector, i.e., it will not be multiplied times the parameters.

ctl.GradCheck scalar, if nonzero and if analytical gradients and/or Hessian have been

provided, numerical gradients and/or Hessian are computed and compared against the analytical versions.

ctl.State	scalar, seed for random number generator.
ctl.Title	string, title of run.
ctl.printIters	scalar, if nonzero, prints iteration information. Default = 0.
ctl.MaxTime	scalar, maximum number of minutes to convergence.
ctl.DisableKey	scalar, if nonzero, keyboard input disabled.

OUTPUT *out* instance of a **optmtResults** structure. For an instance named *out*, the members are:

out.Par	instance of a **PV** structure containing the parameter estimates. Use **pvUnpack** to retrieve matrices and arrays or **pvGetParvector** to get the parameter vector.
out.Fct	scalar, function evaluated at parameters in *out*.**Par**
out.ReturnDescription	string, description of return values.
out.Hessian	K × K matrix, Hessian evaluated at parameters in *out*.**Par**.
out.Gradient	K × 1 vector, gradient evaluated at the parameters in *out*.**Par**.
out.NumIterations	scalar, number of iterations.
out.ElapsedTime	scalar, elapsed time of iterations.
out.Title	string, title of run.
out.Lagrange	K ×2 matrix, Lagrangeans for the bounds constraints.

Whenever a constraint is active, its associated Lagrangean will be nonzero. For any constraint that is inactive throughout the iterations as well as at convergence, the corresponding Lagrangean matrix will be set to a scalar missing value.

out.Retcode return code:

- **0** normal convergence
- **1** forced exit
- **2** maximum number of iterations exceeded
- **3** function calculation failed
- **4** gradient calculation failed
- **5** Hessian calculation failed
- **6** line search failed
- **7** functional evaluation failed
- **8** error with initial gradient
- **10** secant update failed
- **11** maximum time exceeded
- **12** error with weights
- **16** function evaluated as complex
- **20** Hessian failed to invert
- **34** data set could not be opened

REMARKS **Writing the Objective Funcition** There is one required user-provided procedure, the one computing the objective function and optionally the first and/or second derivatives.

The main procedure, computing the objective and optionally the first and/or second derivatives, has three arguments, an instance of type struct **PV** containing the parameters, a second argument that is an

instance of type struct **DS** containing the data, and a third argument that is a vector of zeros and ones indicating which of the results, the function, first derivatives, or second derivatives, are to be computed.

The remaining optional procedures take just two arguments, the instance of the **PV** structure containing the parameters and the the instance of the **DS** structure containing the data.

The instance of the **PV** structure is set up using the **PV** pack procedures, **pvPack**, **pvPackm**, **pvPacks**, and **pvPacksm**. These procedures allow for setting up a parameter vector in a variety of ways.

For example, the following procedure computes the objective and the first derivatives for a tobit model:

```
proc lpr(struct PV p, struct DS d, ind);
    local s2,b0,b,y,x,yh,u,res,g1,g2;

    struct modelResults mm;

    b0 = pvUnpack(p,"b0");
    b = pvUnpack(p,"b");
    s2 = pvUnpack(p,"variance");

    y = d[1].dataMatrix;
    x = d[2].dataMatrix;

    yh = b0 + x * b;
    res = y - yh;
    u = y[.,1] ./= 0;

    if ind[1];
        mm.function = u.*lnpdfmvn(res,s2) +
            (1-u).*(ln(cdfnc(yh/sqrt(s2))));
```

```
    endif;

    if ind[2];
        yh = yh/sqrt(s2);
        g1 = ((res~x.*res)/s2)~((res.*res/s2)-1)/
            (2*s2);
        g2 = ( -( ones(rows(x),1)~x )/
            sqrt(s2) )~(yh/(2*s2));
        g2 = (pdfn(yh)./cdfnc(yh)).*g2;
        mm.gradient = u.*g1 + (1-u).*g2;
    endif;
    retp(mm);

endp;
```

EXAMPLE The following is a complete example. It applys the Biochemical Oxygen Demand model to data taken from Douglas M. Bates and Donald G. Watts, *Nonlinear Regression Analysis and Its Applications*, page 270.

```
library optmt;
#include optmt.sdf

proc lnlk(struct PV p, struct DS d, ind);
    local dev,s2,m,r,b0,b;

    struct modelResults mm;

    b0 = pvUnpack(p,1);
    b = pvUnpack(p,2);
    r = exp(-b*d[2].dataMatrix);
    m = 1 - r;
```

```
    dev = d[1].dataMatrix -  b0*m;
    s2 = dev'dev/rows(dev);

    if ind[1];
        mm.function = lnpdfmvn(dev,s2);
    endif;
    if ind[2];
        mm.gradient =  (dev/s2) .*
  (m ~ b0*d[2].dataMatrix.*r);
    endif;
    retp(mm);

endp;

struct DS d0;
d0 = reshape(dsCreate,2,1);
d0[1].dataMatrix =
  {
     8.3,
    10.3,
    19.0,
    16.0,
    15.6,
    19.8
  };

d0[2].dataMatrix =
  {
    1,
    2,
    3,
    4,
    5,
    7
```

```
    };


struct PV p0;
p0 = pvPacki(pvCreate,19.143,"b0",1);
p0 = pvPacki(p0,.5311,"b",2);


struct optmtControl c0;
c0 = optmtControlCreate;

c0.Bounds = { 10 35,
              0  2 };

struct optmtResults out;
out = optmt(&lnlk,p0,d0,c0);
```

SOURCE optmt.src

OptmtControlCreate

PURPOSE Creates a default instance of type **OptmtControl**.

LIBRARY **optmt**

FORMAT *s* **= OptmtControlCreate;**

OUTPUT *s* instance of type **OptmtControl**.

SOURCE optmtutil.src

OptmtResultsCreate

PURPOSE Creates a default instance of type **OptmtResults**.

LIBRARY **optmt**

FORMAT *s* = **OptmtResultsCreate;**

OUTPUT *s* instance of type **OptmtResults**.

SOURCE optmtutil.src

ModelResultsCreate

PURPOSE Creates a default instance of type **ModelResults**.

LIBRARY **optmt**

FORMAT *s* = **ModelResultsCreate;**

OUTPUT *s* instance of type **ModelResults**.

SOURCE optmtutil.src

OptmtPrt

PURPOSE Formats and prints the output from a call to **Optmt**.

LIBRARY **optmt**

FORMAT *out* = **OptmtPrt(***out***);**

INPUT *out* instance of **optmtResults** structure containing results of an estimation generated by a call to **optmt**.

OUTPUT *out* the input instance of the **optmtResults** structure unchanged.

REMARKS The call to **Optmt** can be nested in the call to **OptmtPrt**:

```
call OptmtPrt(Optmt(&modelProc,par,data,ctl));
```

SOURCE optmtutil.src

Index

H

I

L

M

N

O

P

Q

R

S

T

U

W

X