OUTEST= Output Data Set

The OUTEST= or OUTVAR= output data set saves the optimization solution of PROC NLP. You can use the OUTEST= or OUTVAR= data set as follows:

- to save the values of the objective function on grid points to examine, for example, surface plots using PROC G3D (use the OUTGRID option)
- to avoid any costly computation of analytical (first- or second-order) derivatives during optimization when they are needed only upon termination. In this case a two-step approach is recommended:
    1. In a first execution, the optimization is done; that is, optimal parameter estimates are computed, and the results are saved in an OUTEST= data set.
    2. In a subsequent execution, the optimal parameter estimates in the previous OUTEST= data set are read in an INEST= data set and used with TECH=NONE to compute further results, such as analytical second-order derivatives or some kind of covariance matrix.
- to restart the procedure using parameter estimates as initial values
- to split a time-consuming optimization problem into a series of smaller problems using intermediate results as initial values in subsequent runs. (Refer to the MAXTIME=, MAXIT=, and MAXFUNC= options to trigger stopping.)
- to write the value of the objective function, the parameter estimates, the time in seconds starting at the beginning of the optimization process and (if available) the gradient to the OUTEST= data set during the iterations. After the PROC NLP run is completed, the convergence progress can be inspected by graphically displaying the iterative information. (Refer to the OUTITER option.)

The variables of the OUTEST= data set are

- the BY variables that are used in a DATA= input data set
- a character variable _TECH_ naming the optimization technique used
- a character variable _TYPE_ specifying the type of the observation
- a character variable _NAME_ naming the observation. For a linear constraint, the _NAME_ variable indicates whether the constraint is active at the solution. For the initial observations, the _NAME_ variable indicates if the number in the _RHS_ variable corresponds to the number of positive, negative, or zero eigenvalues.
- numeric variables with the parameter names used in the DECVAR statement. These variables contain a point of the parameter space, lower or upper bound constraints, or the coefficients of linear constraints.
- a numeric variable _RHS_ (right-hand side) that is used for the right-hand-side value of a linear constraint or for the value of the objective function at a point of the parameter space
- a numeric variable _ITER_ that is zero for initial values, equal to the iteration number for the OUTITER output, and missing for the result output

The _TYPE_ variable identifies how to interpret the observation. If _TYPE_ is

- PARMS then parameter-named variables contain the coordinates of the resulting point .

  The _RHS_ variable contains .
- INITIAL then parameter-named variables contain the feasible starting point . The

  RHS variable contains .
- GRIDPNT then (if the OUTGRID option is specified) parameter-named variables contain the

  coordinates of any point used in the grid search. The _RHS_ variable contains .
- GRAD then parameter-named variables contain the gradient at the initial or final estimates.
- STDERR then parameter-named variables contain the approximate standard errors (square roots of the diagonal elements of the covariance matrix) if the COV= option is specified.
- _NOBS_ then (if the COV= option is specified) all parameter variables contain the value of

  NOBS used in computing the value in the formula of the covariance matrix.
- UPPERBD | UB then (if there are boundary constraints) the parameter variables contain the upper bounds.
- LOWERBD | LB then (if there are boundary constraints) the parameter variables contain the lower bounds.
- NACTBC then all parameter variables contain the number of active boundary

  constraints at the solution .
- ACTBC then (if there are active boundary constraints) the observation indicate which parameters are actively constrained, as follows:

  NAME=GE

  the active lower bounds

  NAME=LE

  the active upper bounds

  NAME=EQ

  the active equality constraints
- NACTLC then all parameter variables contain the number of active linear constraints that are recognized as linearly independent.
- NLDACTLC then all parameter variables contain the number of active linear constraints that are recognized as linearly dependent.
- LE then (if there are linear constraints) the observation contains the th linear constraint

  . The parameter variables contain the coefficients , , and the _RHS_ variable contains . If the constraint is active at the solution , then _NAME_=ACTLC or _NAME_=LDACTLC.

- GE then (if there are linear constraints) the observation contains the th linear constraint . The parameter variables contain the coefficients , , and the _RHS_ variable contains . If the constraint is active at the solution , then _NAME_=ACTLC or _NAME_=LDACTLC.
- EQ then (if there are linear constraints) the observation contains the th linear constraint . The parameter variables contain the coefficients , , the _RHS_ variable contains , and _NAME_=ACTLC or _NAME_=LDACTLC.
- LAGRANGE then (if at least one of the linear constraints is an equality constraint or an active inequality constraint) the observation contains the vector of Lagrange multipliers. The Lagrange multipliers of active boundary constraints are listed first followed by those of active linear constraints and those of active nonlinear constraints. Lagrange multipliers are available only for the set of linearly independent active constraints.
- PROJGRAD then (if there are linear constraints) the observation contains the values of the projected gradient in the variables corresponding to the first parameters.
- JACOBIAN then (if the PJACOBI or OUTJAC option is specified) the observations contain the rows of the Jacobian matrix. The _RHS_ variable contains the row number , .
- HESSIAN then the first observations contain the rows of the (symmetric) Hessian matrix. The _RHS_ variable contains the row number , , and the _NAME_ variable contains the corresponding parameter name.
- PROJHESS then the first observations contain the rows of the projected Hessian matrix . The _RHS_ variable contains the row number , , and the _NAME_ variable is blank.
- CRPJAC then the first observations contain the rows of the (symmetric) crossproduct Jacobian matrix at the solution. The _RHS_ variable contains the row number , , and the _NAME_ variable contains the corresponding parameter name.
- PROJCRPJ then the first observations contain the rows of the projected crossproduct Jacobian matrix . The _RHS_ variable contains the row number , , and the _NAME_ variable is blank.

- COV1, COV2, COV3, COV4, COV5, or COV6 then (depending on the COV= option) the first observations contain the rows of the (symmetric) covariance matrix of the parameter estimates. The _RHS_ variable contains the row number , , and the _NAME_ variable contains the corresponding parameter name.
- DETERMIN contains the determinant of the matrix specified by the value of the _NAME_ variable where is the value of the first variable in the DECVAR statement and is in _RHS_.
- NEIGPOS, NEIGNEG, or NEIGZER then the _RHS_ variable contains the number of positive, negative, or zero eigenvalues of the matrix specified by the value of the _NAME_ variable.
- COVRANK then the _RHS_ variable contains the rank of the covariance matrix.
- SIGSQ then the _RHS_ variable contains the scalar factor of the covariance matrix.
- _TIME_ then (if the OUTITER option is specified) the _RHS_ variable contains the number of seconds passed since the start of the optimization.
- TERMINAT then if optimization terminated at a point satisfying one of the termination criteria, an abbreviation of the corresponding criteria is given to the _NAME_ variable. Otherwise _NAME_=PROBLEMS.

If for some reason the procedure does not terminate successfully (for example, no feasible initial values can be computed or the function value or derivatives at the starting point cannot be computed), the OUTEST= data set may contain only part of the observations (usually only the PARMS and GRAD observation).

Note:Generally you can use an OUTEST= data set as an INEST= data set in a further run of PROC NLP. However, be aware that the OUTEST= data set also contains the boundary and general linear constraints specified in the previous run of PROC NLP. When you are using this OUTEST= data set without changes as an INEST= data set, PROC NLP adds the constraints from the data set to the constraints specified by a BOUNDS or LINCON statement. Although PROC NLP automatically eliminates multiple identical constraints you should avoid specifying the same constraint twice.

Output of Profiles

The following observations are written to the OUTEST= data set only when the PROFILE statement or CLPARM option is specified.

Table 6.4 Output of Profiles

TYPE	**_NAME_**	**_RHS_**	**Meaning of Observation**
PLC_LOW	parname	value	coordinates of lower CL for
PLC_UPP	parname	value	coordinates of upper CL for
WALD_CL	LOWER	value	lower Wald CL for in _ALPHA_
WALD_CL	UPPER	value	upper Wald CL for in _ALPHA_

PL_CL	LOWER	value	lower PL CL for in _ALPHA_
PL_CL	UPPER	value	upper PL CL for in _ALPHA_
PROFILE	L(THETA)	missing	value corresponding to in following _NAME_=THETA
PROFILE	THETA	missing	value corresponding to in previous _NAME_=L(THETA)

Assume that the PROFILE statement specifies parameters and confidence levels. For CLPARM, and .

- _TYPE_=PLC_LOW and _TYPE_=PLC_UPP:
  If the CLPARM= option or the PROFILE statement with the OUTTABLE option is specified, then the complete set of parameter estimates (rather than only the confidence limit) is written to the OUTEST= data set for each side of the confidence interval. This output may be helpful for further analyses on how small changes in affect the changes in the other . The _ALPHA_ variable contains the corresponding value of . There should be no more than observations. If the confidence limit cannot be computed, the corresponding observation is not available.
- _TYPE_=WALD_CL:

  If CLPARM=WALD, CLPARM=BOTH, or the PROFILE statement with values is specified, then the Wald confidence limits are written to the OUTEST= data set for each of the default or specified values of . The _ALPHA_ variable contains the corresponding value of . There should be observations.
- _TYPE_=PL_CL:

  If CLPARM=PL, CLPARM=BOTH, or the PROFILE statement with values is specified, then the PL confidence limits are written to the OUTEST= data set for each of the default or specified values of . The _ALPHA_ variable contains the corresponding values of . There should be observations; some observations may have missing values.
- _TYPE_=PROFILE:

  If CLPARM=PL, CLPARM=BOTH, or the CLPARM= statement with or without values is specified, then a set of point coordinates in two adjacent observations with

NAME=L(THETA) (value) and _NAME_=THETA (value) is written to the OUTEST= data set. The _RHS_ and _ALPHA_ variables are not used (are set to missing). The number of observations depends on the difficulty of the optimization problems.