Practical Methods of Optimization

Second Edition

R. Fletcher
Department of Mathmatics and Computer Science
University of Dundee, Scotland, UK

A Wiley-Interscience Publication

JOHN WILEY & SONS
Chichester · New York · Brisbane · Toronto · Singapore

Contents

Table of Notation

$\mathbf{A}$	matrix (Jacobian matrix, matrix of constraint normals)
$\mathbf{I}$	unit matrix
$\mathbf{L}, \mathbf{U}$	lower or upper triangular matrices respectively
$\mathbf{P}, \mathbf{Q}$	permutation matrix or orthogonal matrix
$\mathbf{a}$	vector (usually a column vector)
$\mathbf{a}_i\ i = 1, 2, \ldots$	set of vectors (columns of $\mathbf{A}$)
$\mathbf{e}_i\ i = 1, 2, \ldots$	coordinate vectors (columns of $\mathbf{I}$)
$\mathbf{e}$	vector of ones $(1, 1, \ldots, 1)^T$
$\mathbf{A}^T, \mathbf{a}^T$	transpose
$\mathbb{R}^n$	n-dimensional space
$\mathbf{x}$	variables in an optimization problem
$\mathbf{x}^{(k)}\ k = 1, 2, \ldots$	iterates in an iterative method
$\mathbf{x}^*$	local minimizer or local solution
$\mathbf{s}, \mathbf{s}^{(k)}$	search direction (on iteration k)
$\alpha, \alpha^{(k)}$	step length (on iteration k)
$\boldsymbol{\delta}, \boldsymbol{\delta}^{(k)}$	correction to $\mathbf{x}^{(k)}$
$f(\mathbf{x})$	objective function
∇	first derivative operator (elements $\partial/\partial x_i$)
$\mathbf{g}(\mathbf{x}) = \nabla f(\mathbf{x})$	gradient vector
$f^*, \mathbf{g}^*, \ldots$	$f(\mathbf{x}^*), \mathbf{g}(\mathbf{x}^*), \ldots$
$f^{(k)}, \mathbf{g}^{(k)}, \ldots$	$f(\mathbf{x}^{(k)}), \mathbf{g}(\mathbf{x}^{(k)}), \ldots$
∇^2	second derivative operator (elements $\partial^2/\partial x_i \partial x_j$)
$\mathbf{G}(\mathbf{x}) = \nabla^2 f(\mathbf{x})$	Hessian matrix (second derivative matrix)
$\mathbb{C}^k$	set of k times continuously differentiable functions
$l(\mathbf{x})$	linear function (1.2.8)
$q(\mathbf{x})$	quadratic function (1.2.11)
$[a, b]$	closed interval
(a, b)	open interval
$\lVert \cdot \rVert$	norm of a vector or matrix
$\square$	end of proof
$\exists, \forall$	'there exists', 'for all'
$\Rightarrow, \Leftrightarrow$	'implies', 'equivalent to'
$O(.), o(.)$	'big O' and 'little o' notation (Hardy, 1960) (let $h \to 0$: then $a = O(h)$ iff $\exists$ a constant c such that $\lvert a \rvert \leqslant ch$, and $a = o(h)$ iff $a/h \to 0$)
$\subset$	set inclusion
$\in$	element in set

$\emptyset$	empty set
$\triangleq$	equal by definition
$\mathscr{L}$	Lagrangian function
σ, ρ, τ	fixed parameters in algorithms
$\mathbf{r}$	residual vector, artificial variables vector
$\mathbf{c}_i(\mathbf{x})\ i = 1, 2, \ldots$	constraint functions
E, I	set of equality and inequality constraints respectively (I is set of integer variables in Chapter 13)
$\mathscr{A}$	set of active constraints
$\mathbf{c}(\mathbf{x})$	vector of constraint functions (usually for equality constraints only)
$\mathbf{a}_i(\mathbf{x}) = \nabla \mathbf{c}_i(\mathbf{x})$	constraint gradient vector (normal vector)
x_i^+, x_i^-	$\max(x, 0)$ and $\max(-x, 0)$ respectively
$\mathbf{A}_B$	basis matrix in linear programming
B, N	sets of basic and nonbasic variables respectively
$\hat{\mathbf{A}}, \hat{\mathbf{c}}, \hat{\mathbf{b}}$	tableau matrix, reduced costs, basic variable values
$\mathbf{x}_c, \mathbf{x}_+$	current and next point (Section 8.7)
$\lambda_i\ i = 1, 2, \ldots$	Lagrange multipliers
$\boldsymbol{\lambda}$	vector of Lagrange multipliers (usually for equality constraints only)
I^*	$\mathscr{A}^* \cap \mathrm{I}$ (set of active inequality constraints at $\mathbf{x}^*$)
$\mathscr{F}, F, \mathscr{G}, G$	feasible direction sets for optimality conditions
K	convex set
$\mathbf{Y}, \mathbf{Z}$	left inverse and null space matrices in generalized elimination
$:=$	assignment operator
$q^{(k)}(\boldsymbol{\delta})$	model quadratic function obtained by Taylor series expansion about $\mathbf{x}^{(k)}$
$\mathbf{l}^{(k)}(\boldsymbol{\delta})$	linear model about $\mathbf{x}^{(k)}$ to vector of constraint functions
$\|\mathscr{A}\|$	number of elements in active set
ν, σ	weighting parameters in penalty functions
$\dot{\mathbf{x}}$	$\mathrm{d}x/\mathrm{d}\theta$ for some trajectory $\mathbf{x}(\theta)$
∇_x, ∇_λ	partial derivative operators with respect to x, λ respectively
$\boldsymbol{\nabla} = \begin{pmatrix} \nabla_x \\ \nabla_\lambda \end{pmatrix}$	combined vector of partial derivatives
$\mathbf{W} = \nabla_x^2 \mathscr{L}(\mathbf{x}, \boldsymbol{\lambda})$	Hessian of Lagrangian function
$\mathbf{W}^*, \mathbf{W}^{(k)}, \ldots$	$\mathbf{W}(\mathbf{x}^*, \boldsymbol{\lambda}^*), \mathbf{W}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)}), \ldots$
$\mathbf{M}$	approximation to reduced Hessian matrix ($\mathbf{Z}^T\mathbf{G}\mathbf{Z}$ in Chapter 11, or $\mathbf{Z}^T\mathbf{W}\mathbf{Z}$ in Chapter 12)
$\phi(\mathbf{x})$	penalty function (possibly with additional parameters in argument list)
$\psi^{(k)}(\boldsymbol{\delta})$	model approximating function about $\mathbf{x}^{(k)}$ to $\phi(\mathbf{x}^{(k)} + \boldsymbol{\delta})$
$[\mathrm{x}]$	greatest integer not larger than x
G, T	graph, tree
$h(\mathbf{c})$	convex non-smooth function
$\partial h(\mathbf{c})$	subdifferential (set of all subgradients at $\mathbf{c}$)
$h^*, h^{(k)}, \ldots$	$h(\mathbf{c}^*), h(\mathbf{c}^{(k)}), \ldots$ (may be $h(\mathbf{c}(\mathbf{x}^*))$ etc. when $\mathbf{c} = \mathbf{c}(\mathbf{x})$)
$\mathbf{D}^*$	basis vectors for subdifferential (14.2.30)
$\partial h - \boldsymbol{\lambda}$	the set $\{\mathbf{u} : \mathbf{u} = \boldsymbol{\gamma} - \boldsymbol{\lambda}, \boldsymbol{\gamma} \in \partial h\}$
$\partial h \backslash \boldsymbol{\lambda}$	the set $\{\mathbf{u} : \mathbf{u} \in \partial h,\ \mathbf{u} \neq \boldsymbol{\lambda}\}$

References

Abadie, J. and Carpentier, J. (1969). Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints, in *Optimization* (Ed. R. Fletcher), Academic Press, London.

Akaike, H. (1959). On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method, *Ann. Inst. Stat. Math. Tokyo*, **11**, 1–16.

Al-Baali, M. (1985). Descent property and global convergence of the Fletcher–Reeves method with inexact line search, *IMA J. Num. Anal.*, **5**, 121–124.

Al-Baali, M. and Fletcher, R. (1985). Variational methods for nonlinear least squares, *J. Oper. Res. Soc.*, **36**, 405–421.

Al-Baali, M. and Fletcher, R. (1986). An efficient line search for nonlinear least squares, *J. Opt. Theo. Applns.*, **48**, 359–378.

Andreassen, D. O. and Watson, G. (1976). Nonlinear Chebyshev approximation subject to constraints, *J. Approx. Theo.*, **18**, 241–250.

Apostol, T. M. (1957). *Mathematical Analysis*, Addison-Wesley, Reading, Mass.

Appelgren, L. (1971). Integer programming methods for a vessel scheduling problem, *Transportation Science*, **5**, 64–78.

Balinski, M. L. and Gomory, R. E. (1963). A mutual primal-dual simplex method, in *Recent Advances in Mathematical Programming* (Eds R. L. Graves and P. Wolfe), McGraw-Hill, New York.

Balinski, M. L. and Wolfe, P. (Eds) (1975). *Nondifferentiable Optimization*, Mathematical Programming Study 3, North-Holland, Amsterdam.

Bandler, J. W. and Charalambous, C. (1972). Practical least p-th optimization of networks, *IEEE Trans. Microwave Theo. Tech., (1972 Symposium Issue)*, **20**, 834–840.

Barnes, J. G. P. (1965). An algorithm for solving nonlinear equations based on the secant method, *Computer J.*, **8**, 66–72.

Barrodale, I. (1970). On computing best L_1 approximations, in *Approximation Theory* (Ed. A. Talbot), Academic Press, London.

Barrodale, I. and Roberts, F. D. K. (1973). An improved algorithm for discrete l_1 linear approximation, *SIAM J. Num. Anal.*, **10**, 839–848.

Bartels, R. H. (1971). A stabilization of the simplex method, *Numer. Math.*, **16**, 414–434.

Beale, E. M. L. (1959). On quadratic programming, *Naval Res. Log. Quart.*, **6**, 227–244.

Beale, E. M. L. (1967). Numerical methods, in *Nonlinear Programming* (Ed. J. Abadie), North-Holland, Amsterdam.

Beale, E. M. L. (1968). *Mathematical Programming in Practice*, Pitman, London.

Beale, E. M. L. (1970). Advanced algorithmic features for general mathematical programming systems, in *Integer and Nonlinear Programming* (Ed. J. Abadie), North-Holland, Amsterdam.

Beale, E. M. L. (1972). A derivation of conjugate gradients, in *Numerical Methods for Nonlinear Optimization* (Ed. F. A. Lootsma), Academic Press, London.

Beale, E. M. L. (1978). Integer programming, in *The State of the Art in Numerical Analysis* (Ed. D. A. H. Jacobs), Academic Press, London.

Bennett, J. M. (1965). Triangular factors of modified matrices, *Numer. Math.*, **7**, 217–221.

Benveniste, R. (1979). A quadratic programming algorithm using conjugate search directions, *Math. Prog.*, **16**, 63–80.

Biggs, M. C. (1973). A note on minimization algorithms which make use of non-quadratic properties of the objective function, *J. Inst Math. Applns.*, **12**, 337–338.

Biggs, M. C. (1975). Constrained minimization using recursive quadratic programming: some alternative subproblem formulations, in *Towards Global Optimization* (Eds L. C. W. Dixon and G. P. Szego), North-Holland, Amsterdam.

Biggs, M. C. (1978). On the convergence of some constrained minimization algorithms based on recursive quadratic programming, *J. Inst. Maths. Applns.*, **21**, 67–81.

Björck, A. (1985). Stability analysis of the method of semi-normal-equations for linear least squares problems, Linköping Univ. Report LiTH-MAT-R-1985–08.

Boggs, P. T. and Tolle, J. W. (1984). A family of descent functions for constrained optimization, *SIAM J. Num. Anal.*, **21**, 1146–1161.

Boggs, P. T., Tolle, J. W. and Wang, P. (1982). On the local convergence of quasi-Newton methods for constrained optimization, *SIAM J. Control Optim.*, **20**, 161–171.

Bradley, J. and Clyne, H. M. (1976). Applications of geometric programming to building design problems, in *Optimization in Action* (Ed. L. C. W. Dixon), Academic Press, London.

Branin, F. H. and Hoo, S. K. (1972). A method for finding multiple extrema of a function of n variables, in *Numerical Methods for Nonlinear Optimization* (Ed. F. A. Lootsma), Academic Press, London.

Brent, R. P. (1973a). *Algorithms for Minimization without Derivatives*, Prentice-Hall Inc., Englewood Cliffs, N. J.

Brent, R. P. (1973b). Some efficient algorithms for solving systems of nonlinear equations, *SIAM J. Num. Anal.*, **10**, 327–344.

Breu, R. and Burdet, C.-A. (1974). Branch and bound experiments in zero-one programming, in *Approaches to Integer Programming* (Ed. M. L. Balinski), Mathematical Programming Study 2, North-Holland, Amsterdam.

Brodlie, K. W. (1975). A new direction set method for unconstrained minimization without evaluating derivatives, *J. Inst. Maths. Applns.*, **15**, 385–396.

Brodlie, K. W. (1977). Unconstrained minimization, in *The State of the Art in Numerical Analysis* (Ed. D. A. H. Jacobs), Academic Press, London.

Brodlie, K. W., Gourlay, A. R. and Greenstadt, J. L. (1973). Rank-one and rank-two corrections to positive definite matrices expressed in product form, *J. Inst. Maths. Applns.*, **11**, 73–82.

Brown, K. M. and Dennis, J. E. (1971). A new algorithm for nonlinear least squares curve fitting, in *Mathematical Software* (Ed. J. R. Rice), Academic Press, New York.

Broyden, C. G. (1965). A class of methods for solving nonlinear simultaneous equations, *Maths. Comp.*, **19**, 577–593.

Broyden, C. G. (1967). Quasi-Newton methods and their application to function minimization, *Maths. Comp.*, **21**, 368–381.

Broyden, C. G. (1970). The convergence of a class of double rank minimization algorithms, parts I and II, *J. Inst. Maths. Applns.*, **6**, 76–90 and 222–231.

Broyden, C. G. (1975). *Basic Matrices*, Macmillan, London.

Broyden, C. G., Dennis, J. E. and Moré, J. J. (1973). On the local and superlinear convergence of quasi-Newton methods, *J. Inst. Maths. Applns.*, **12**, 223–245.

Buckley, A. (1975). Constrained minimization using Powell's conjugacy approach, AERE Harwell report CSS 22.

Bunch, J. R. and Parlett, B. N. (1971). Direct methods for solving symmetric indefinite systems of linear equations, *SIAM J. Num. Anal.*, **8**, 639–655.

Buys, J. D. (1972). Dual algorithms for constrained optimization problems, Ph.D. thesis, Univ. of Leiden.

Byrd, R. H. (1984). On the convergence of constrained optimization methods with accurate Hessian information on a subspace, Univ. of Colorado at Boulder, Dept. of Comp. Sci. Report CU-CS-270-84.

Byrd, R. H. and Shultz, G. A. (1982). A practical class of globally convergent active set strategies for linearly constrained optimization, Univ. of Colorado at Boulder, Dept. of Comp. Sci. Report CU-CS-238-82.

Carroll, C. W. (1961). The created response surface technique for optimizing nonlinear restrained systems, *Operations Res.*, **9**, 169–84.

Chamberlain, R. M. (1979). Some examples of cycling in variable metric methods for constrained minimization, *Math. Prog.*, **16**, 378–383.

Chamberlain, R. M., Lemarechal, C., Pedersen, H. C. and Powell, M. J. D. (1982). The watchdog technique for forcing convergence in algorithms for constrained optimization, in *Algorithms for Constrained Minimization of Smooth Nonlinear Functions* (Eds A. G. Buckley and J.-L. Goffin), Mathematical Programming Study 16, North-Holland, Amsterdam.

Charalambous, C. (1977). Nonlinear least p-th optimization and nonlinear programming, *Math. Prog.*, **12**, 195–225.

Charalambous, C. (1979). Acceleration of the least p-th algorithm for minimax optimization with engineering applications, *Math. Prog.*, **17**, 270–297.

Charalambous, C. and Conn, A. R. (1978). An efficient method to solve the minimax problem directly, *SIAM J. Num. Anal.*, **15**, 162–187.

Charnes, A. (1952). Optimality and degeneracy in linear programming, *Econometrica*, **20**, 160–170.

Clarke, F. H. (1975). Generalized gradients and applications, *Trans. Amer. Math. Soc.*, **205**, 247–262.

Coleman, T. F. and Conn, A. R. (1982a). Nonlinear programming via an exact penalty function: asymptotic analysis, *Math. Prog.*, **24**, 123–136.

Coleman, T. F. and Conn, A. R. (1982b). Nonlinear programming via an exact penalty function: global analysis, *Math. Prog.*, **24**, 137–161.

Colville, A. R. (1968). A comparative study on nonlinear programming codes, IBM NY Scientific Center Report 320–2949.

Concus, P. (1967). Numerical solution of the minimal surface equation, *Maths. Comp.*, **21**, 340–350.

Conn, A. R. (1973). Constrained optimization using a nondifferentiable penalty function, *SIAM J. Num. Anal.*, **13**, 145–154.

Conn, A. R. (1979). An efficient second order method to solve the (constrained) minimax problem, Univ. of Waterloo Dept. of Combinatorics and Optimization Res. Report CORR-79-5.

Conn, A. R. and Sinclair, J. W. (1975). Quadratic programming via a non-differentiable penalty function, Univ. of Waterloo Dept. of Combinatorics and Optimization Report CORR 75/15.

Coope, I. D. (1976). Conjugate direction algorithms for unconstrained optimization, Ph.D. thesis, Univ. of Leeds Dept. of Mathematics.

Coope, I. D. and Fletcher, R. (1980). Some numerical experience with a globally convergent algorithm for nonlinearly constrained optimization, *J. Opt. Theo. Applns.*, **32**, 1–16.

Cottle, R. W. and Dantzig, G. B. (1968). Complementary pivot theory of mathematical programming, *J. Linear Algebra Applns.*, **1**, 103–125.

Courant, R. (1943). Variational methods for the solution of problems of equilibrium and vibration, *Bull. Amer. Math. Soc.*, **49**, 1–23.

Curtis, A. R., Powell, M. J. D. and Reid, J. K. (1974). On the estimation of sparse Jacobian matrices, *J. Inst. Maths. Applns.*, **13**, 117–120.

Curry, H. (1944). The method of steepest descent for nonlinear minimization problems, *Quart. Appl. Math.*, **2**, 258–261.

Dantzig, G. B. (1963). *Linear Programming and Extensions*, Princeton University Press, Princeton, N. J.

Dantzig, G. B., Orden, A. and Wolfe, P. (1955). The generalized simplex method for minimizing a linear form under linear inequality restraints, *Pacific J. Maths.*, **5**, 183–195.

Dantzig, G. B. and Wolfe, P. (1960). A decomposition principle for linear programs, *Operations Res.*, **8**, 101–111.

Davidon, W. C. (1959). Variable metric method for minimization, AEC Res. and Dev. Report ANL-5990 (revised).

Davidon, W. C. (1968). Variance algorithms for minimization, *Computer J.* **10**, 406–410.

Davidon, W. C. (1975). Optimally conditioned optimization algorithms without line searches, *Math. Prog.*, **9**, 1–30.

Davidon, W. C. (1980). Conic approximations and collinear scaling for optimizers, *SIAM J. Num. Anal.*, **17**, 268–281.

Dembo, R. S. (1976). A set of geometric programming test problems and their solutions, *Math. Prog.*, **10**, 192–213.

Dembo, R. S. (1979). Second order algorithms for the posynomial geometric programming dual, Part I: Analysis, *Math. Prog.*, **17**, 156–175.

Dembo, R. S. and Avriel, M. (1978). Optimal design of a membrane separation process using geometric programming, *Math. Prog.*, **15**, 12–25.

Dembo, R. S. and Klincewicz, J. G. (1981). A scaled reduced gradient algorithm for network flow problems with convex separable costs, in *Network Models and Applications* (Ed. D. Klingman and J. M. Mulvey), Mathematical Programming Study 15, North-Holland Amsterdam.

Demyanov, V. F. and Malozemov, V. N. (1971). The theory of nonlinear minimax problems, *Uspekhi Matematcheski Nauk*, **26**, 53–104.

Dennis, J. E. (1977). Nonlinear least squares and equations, in *The State of the Art in Numerical Analysis* (Ed. D. A. H. Jacobs), Academic Press, London.

Dennis, J. E., Gay, D. M. and Welsch, R. E. (1981). An adaptive nonlinear least-squares algorithm, *ACM Trans. Math. Software*, **7**, 348–368

Di Pillo, G. and Grippo, L. (1979). A new class of augmented Lagrangians in nonlinear programming, *SIAM J. Control Optim.*, **17**, 618–628.

Dixon, L. C. W. (1972). Quasi-Newton algorithms generate identical points, *Math. Prog.*, **2**, 383–387.

Dixon, L. C. W. (1973). Conjugate directions without line searches, *J. Inst. Maths. Applns.*, **11**, 317–328.

Dixon, L. C. W. (ed.) (1976). *Optimization in Action*, Academic Press, London.

Dixon, L. C. W. and Szego, G. P. (eds) (1975). *Towards Global Optimization*, North-Holland, Amsterdam.

Duffin, R. J., Peterson, E. L. and Zener, C. (1967). *Geometric Programming—Theory and Application*, John Wiley, New York.

El-Attar, R. A., Vidyasagar, M. and Dutta, S. R. K. (1979). An algorithm for l_1-norm minimization with application to nonlinear l_1 approximation, *SIAM J. Num. Anal.*, **16**, 70–86.

Eriksson, J. (1980). A note on solution of large sparse maximum entropy problems with linear equality constraints, *Math. Prog.*, **18**, 146–154.

Fiacco, A. V. and McCormick, G. P. (1968). *Nonlinear Programming*, John Wiley, New York.

Fisher, M. L., Northup, W. D. and Shapiro, J. F. (1975). Using duality to solve discrete optimization problems: theory and computational experience, in *Nondifferentiable Optimization* (Eds M. L. Balinski and P. Wolfe), Mathematical Programming Study 3, North-Holland, Amsterdam.

Fletcher, R. (1965). Function minimization without evaluating derivatives—a reivew, *Computer J.*, **8**, 33–41.

Fletcher, R. (1969). A technique for orthogonalization, *J. Inst. Maths. Applns.*, **5**, 162–116.

Fletcher, R. (1970a). A new approach to variable metric algorithms, *Computer J.*, **13**, 317–322.

Fletcher, R. (1970b). The calculation of feasible points for linearly constrained optimization problems, AERE Harwell Report AERE-R6354.

Fletcher, R. (1971a). A modified Marquardt subroutine for nonlinear least squares, AERE Harwell Report AERE-R6799.

Fletcher, R. (1971b). A general quadratic programming algorithm, *J. Inst. Maths. Applns.*, **7**, 76–91.

Fletcher, R. (1972a). Conjugate direction methods, in *Numerical Methods for Unconstrained Optimization* (Ed. W. Murray), Academic Press, London.

Fletcher, R. (1972b). An algorithm for solving linearly constrained optimization problems, *Math. Prog.*, **2**, 133–165.

Fletcher, R. (1972c). Minimizing general functions subject to linear constraints, in *Numerical Methods for Nonlinear Optimization* (Ed. F. A. Lootsma), Academic Press, London.

Fletcher, R. (1973). An exact penalty function for nonlinear programming with inequalities, *Math. Prog.*, **5**, 129–150.

Fletcher, R. (1975). An ideal penalty function for constrained optimization, *J. Inst. Maths. Applns.*, **15**, 319–342, and in *Nonlinear Programming 2* (Eds O. L. Mangasarian, R. R. Meyer and S. M. Robinson), Academic Press, London (1975).

Fletcher, R. (1978). On Newton's method for minimization, in *Proc. IX Int. Symp. on Math. Programming* (Ed. A. Prekopa), Akademiai Kiado, Budapest.

Fletcher, R. (1981). Numerical experiments with an L_1 exact penalty function method, in *Nonlinear Programming 4* (Eds O. L. Mangasarian, R. R. Meyer and S. M. Robinson), Academic Press, New York.

Fletcher, R. (1982a). A model algorithm for composite nondifferentiable optimization problems, in *Nondifferential and Variational Techniques in Optimization* (Eds D. C. Sorensen and R. J.-B. Wets), Mathematical Programming Study 17, North-Holland, Amsterdam.

Fletcher, R. (1982b). Second order corrections for nondifferentiable optimization, in *Numerical Analysis, Dundee 1981* (Ed. G. A. Watson), Lecture Notes in Mathematics 912, Springer-Verlag, Berlin.

Fletcher, R. (1985a). Degeneracy in the presence of round-off errors, Univ. of Dundee Dept. of Math. Sci. Report NA/79.

Fletcher, R. (1985b). An l_1 penalty method for nonlinear constraints, in *Numerical Optimization 1984* (Eds P. T. Boggs, R. H. Byrd and R. B. Schnabel), SIAM Publications, Philadelphia.

Fletcher, R. and Freeman, T. L. (1977). A modified Newton method for minimization, *J. Opt. Theo. Applns.*, **23**, 357–372.

Fletcher, R. and Jackson, M. P. (1974). Minimization of a quadratic function of many variables subject only to lower and upper bounds, *J. Inst. Maths. Applns.*, **14**, 159–174.

Fletcher, R. and Matthews, S. P. J. (1984). Stable modification of explicit LU factors for simplex updates, *Math. Prog.*, **30**, 267–284.

Fletcher, R. and Matthews, S. P. J. (1985). A stable algorithm for updating triangular factors under a rank one change, *Maths. Comp.*, **45**, 471–485.

Fletcher, R. and McCann, A. P. (1969). Acceleration techniques for nonlinear programming, in *Optimization* (Ed. R. Fletcher), Academic Press, London.

Fletcher, R. and Powell, M. J. D. (1963). A rapidly convergent descent method for minimization, *Computer. J.*, **6**, 163–168.

Fletcher, R. and Powell, M. J. D. (1974). On the modification of LDL^T factorizations, *Maths. Comp.*, **29**, 1067–1087.

Fletcher, R. and Reeves, C. M. (1964). Function minimization by conjugate gradients, *Computer J.*, **7**, 149–154.

Fletcher, R. and Sainz de la Maza, E. (1987). Nonlinear programming and nonsmooth optimization by successive linear programming, Univ. of Dundee, Dept. of Math. Sci. Report NA/100.

Fletcher, R. and Sinclair, J. W. (1981). Degenerate values for Broyden methods, *J. Opt. Theo. Applns.*, **33**, 311–324.

Fletcher, R. and Watson, G. A. (1980). First and second order conditions for a class of nondifferentiable optimization problems, *Math. Prog.*, **18**, 291–307; abridged from a Univ. of Dundee Dept. of Mathematics Report NA/28 (1978).

Fletcher, R. and Xu, C. (1985). Hybrid methods for nonlinear least squares, Univ. of Dundee Dept. of Math. Sci. Report NA/92, (to appear in *IMA J. Num. Anal.*).

Forrest, J. J. H. and Tomlin, J. A. (1972). Updated triangular factors of the basis to maintain sparsity in the product form simplex method, *Math. Prog.*, **2**, 263–278.

Frisch, K. R. (1955). The logarithmic potential method of convex programming, Oslo Univ. Inst. of Economics Memorandum, May 1955.

Gacs, P. and Lovasz, L. (1979). Khachian's algorithm for linear programming, Stanford Univ. Dept. of Comp. Sci. Report CS 750.

Gay, D. M. (1979). Some convergence properties of Broyden's method, *SIAM J. Num. Anal.*, **16**, 623–630.

Gay, D. M. (1985). A variant of Karmarkar's linear programming algorithm for problems in standard form, AT&T Bell Labs, Num. Anal. Manuscript 85–10.

Gentleman, W. M. (1973). Least squares computations by Givens' transformations without square roots, *J. Inst. Maths. Applns.*, **12**, 329–336.

Gerber, R. R. and Luk, F. T. (1980). A generalized Broyden's method for solving simultaneous linear equations, Cornell Univ. Dept. of Comp. Sci. Report TR-80-438.

Gill, P. E. and Murray, W. (1972). Quasi-Newton methods for unconstrained optimization, *J. Inst. Maths. Applns.*, **9**, 91–108.

Gill, P. E. and Murray, W. (1973). A numerically stable form of the simplex algorithm, *J. Linear Algebra Applns.*, **7**, 99–138.

Gill, P. E. and Murray, W. (1974a). Newton type methods for linearly constrained optimization, in *Numerical Methods for Constrained Optimization* (Eds P. E. Gill and W. Murray), Academic Press, London.

Gill, P. E. and Murray, W. (1974b). Methods for large-scale linearly constrained problems, in *Numerical Methods for Constrained Optimization* (Eds P. E. Gill and W. Murray), Academic Press, London.

Gill, P. E. and Murray, W. (1974c). Quasi-Newton methods for linearly constrained optimization, in *Numerical Methods for Constrained Optimization* (Eds P. E. Gill and W. Murray), Academic Press, London.

Gill, P. E. and Murray, W. (1976a). Nonlinear least squares and nonlinearly constrained optimization, in *Numerical Analysis, Dundee 1975* (Ed. G. A. Watson), Lecture Notes in Mathematics 506, Springer-Verlag, Berlin.

Gill, P. E. and Murray, W. (1976b). Minimization of a nonlinear function subject to bounds on the variables, NPL Report NAC 72.

Gill, P. E. and Murray, W. (1978a). Modification of matrix factorizations after a rank-one change, in *The State of the Art in Numerical Analysis* (Ed. D. A. H. Jacobs), Academic Press, London.

Gill, P. E. and Murray, W. (1978b). Numerically stable methods for quadratic programming, *Math. Prog.*, **14**, 349–372.
Gill, P. E., Murray, W. and Picken, S. M. (1972). The implementation of two modified Newton algorithms for unconstrained optimization, NPL Report NAC24.
Gill, P. E., Murray, W. and Pitfield, R. A. (1972). The implementation of two revised quasi-Newton algorithms for unconstrained optimization, NPL Report NAC11.
Gill, P. E., Gould, N. I. M., Murray, W., Saunders, M. A. and Wright, M. H. (1984a). Weighted Gram–Schmidt method for convex quadratic programming, *Math. Prog.*, **30**, 176–195.
Gill, P. E., Murray, W., Saunders, M. A. and Wright, M. H. (1984b). Sparse matrix methods in optimization, *SIAM J. Sci. Stat. Comp.*, **5**, 562–589.
Gill, P. E., Murray, W., Saunders, M. A., Tomlin, J. A. and Wright, M. H. (1985). On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method, Stanford Univ. Systems Optimization Lab. Report SOL 85–11.
Gill, P. E., Murray, W., Saunders, M. A. and Wright, M. H. (1986). Some theoretical properties of an augmented Lagrangian merit function, Stanford Univ. Systems Optimization Lab. Report SOL 86–6.
Goldfarb, D. (1969). Extension of Davidon's variable metric method to maximization under linear inequality and equality constraints, *SIAM J. Appl. Math.*, **17**, 739–764.
Goldfarb, D. (1970). A family of variable metric methods derived by variational means, *Maths. Comp.*, **24**, 23–26.
Goldfarb, D. (1972). Extensions of Newton's method and simplex methods for solving quadratic programs, in *Numerical Methods for Nonlinear Optimization* (Ed. F. A. Lootsma), Academic Press, London.
Goldfarb, D. (1980). Curvilinear path steplength algorithms for minimization which use directions of negative curvature, *Math. Prog.*, **18**, 31–40.
Goldfarb, D. and Idnani, A. (1981). Dual and primal-dual methods for solving strictly convex quadratic programs, in *IIMAS Workshop in Numerical Analysis (1981) Mexico* (Ed. J. P. Hennart), Lecture Notes in Mathematics, Springer-Verlag, Berlin.
Goldfarb, D. and Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs, *Math. Prog.*, **27**, 1–33.
Goldfarb, D. and Reid, J. K. (1977). A practicable steepest-edge algorithm, *Math. Prog.*, **12**, 361–371.
Goldfeld, S. M., Quandt, R. E. and Trotter, H. F. (1966). Maximisation by quadratic hill-climbing, *Econometrica*, **34**, 541–551.
Goldstein, A. A. (1965). On steepest descent, *SIAM J. Control*, **3**, 147–151.
Goldstein, A. A. and Price, J. F. (1967). An effective algorithm for minimization, *Numer. Math.*, **10**, 184–189.
Grandinetti, L. (1979). Factorization versus nonfactorization in quasi-Newtonian algorithms for differentiable optimization, in *Methods for Operations Research*, Hain Verlay.
Graves, G. W. and Brown, G. G. (1979). Computational implications of degeneracy in large scale mathematical programming, X Symposium on Mathematical Programming, Montreal, August 1979.
Greenstadt, J. L. (1970). Variations of variable metric methods, *Maths. Comp.*, **24**, 1–22.
Grigoriadis, M. (1982). Minimum-cost network flows, Part I: an implementation of the network simplex method, Rutgers Univ. Lab. for Comp. Sci. Report LCSR-TR-37.
Griewank, A. and Toint, Ph. L. (1984). Numerical experiments with partially separable optimization problems, in *Numerical Analysis, Dundee 1983* (Ed. D. F. Griffiths), Lecture Notes in Mathematics 1066, Springer-Verlag, Berlin.
Hadley, G. (1961). *Linear Algebra*, Addison-Wesley, Reading, Mass.
Hadley, G. (1962). *Linear Programming*, Addison-Wesley, Reading, Mass.

Hald, J. and Madsen, K. (1981). Combined LP and quasi-Newton methods for minimax optimization, *Math. Prog.*, **20**, 49–62.

Hald, J. and Madsen, K. (1985). Combined LP and quasi-Newton methods for nonlinear l_1 optimization, *SIAM J. Num. Anal.*, **22**, 68–80.

Han, S. P. (1976). Superlinearly convergent variable metric algorithms for general nonlinear programming problems, *Math. Prog.*, **11**, 263–282.

Han, S. P. (1977). A globally convergent method for nonlinear programming, *J. Opt. Theo. Applns.*, **22**, 297–309.

Han, S. P. and Mangasarian, O. L. (1979). Exact penalty functions in nonlinear programming, *Math. Prog.*, **17**, 251–269.

Hardy, G. H. (1960). *A Course of Pure Mathematics* (10th edn), Cambridge Univ. Press, Cambridge, England.

Hebden, M. D. (1973). An algorithm for minimization using exact second derivatives, AERE Harwell Report TP515.

Hestenes, M. R. (1969). Multiplier and gradient methods, *J. Opt. Theo. Applns.*, **4**, 303–320, and in *Computing Methods in Optimization Problems, 2* (Eds L. A. Zadeh, L. W. Neustadt and A. V. Balakrishnan), Academic Press, New York (1969).

Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems, *J. Res. N.B.S.*, **49**, 409–436.

Holt, J. N. and Fletcher, R. (1979). An algorithm for constrained nonlinear least squares, *J. Inst. Maths. Applns.*, **23**, 449–464.

Hoshino, S. (1972). A formulation of variable metric methods, *J. Inst. Maths. Applns.*, **10**, 394–403.

Huang, H. Y. (1970). Unified approach to quadratically convergent algorithms for function minimization, *J. Opt. Theo. Applns.*, **5**, 405–423.

Jensen, P. A. and Barnes, J. W. (1980). *Network Flow Programming*, John Wiley, New York.

Kamesam, P. V. and Meyer, R. R. (1984). Multipoint methods for separable nonlinear networks, in *Mathematical Programming at Oberwolfach* (Eds B. Korte and K. Ritter), Mathematical Programming Study 22, North-Holland, Amsterdam.

Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming, *Combinatorics*, **4**, 373–395.

Kennington, J. L. and Helgason, R. V. (1980). *Algorithms for Network Programming*, John Wiley, New York.

Kershaw, D. S. (1978). The incomplete Cholesky-conjugate gradient method for the iterative solution of systems of linear equations, *J. Comp. Phys.*, **26**, 43–65.

Khachiyan, L. G. (1979). A polynomial algorithm in linear programming, *Doklady Akad. Nauk. USSR*, **244**, 1093–1096, translated as *Soviet Mathematics Doklady*, **20**, 191–194.

Kiefer, J. (1957). Optimal sequential search and approximation methods under minimum regularity conditions, *SIAM J. Appl. Math.*, **5**, 105–136.

Klee, V. and Minty, G. J. (1971). How good is the simplex algorithm?, in *Inequalities III* (Ed. O. Shisha), Academic Press.

Kuhn, H. W. and Tucker, A. W. (1951). Nonlinear programming, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Ed. J. Neyman), University of California Press.

Lancaster, P. (1969). *Theory of Matrices*, Academic Press, New York.

Lasdon, L. S. (1985). Nonlinear programming algorithms—applications, software and comparisons, in *Numerical Optimization 1984* (Eds P. T. Boggs, R. H. Byrd and R. B. Schnabel), SIAM Publications, Philadelphia.

Lemarechal, C. (1978). Bundle methods in nonsmooth optimization, in *Nonsmooth Optimization* (Eds C. Lemarechal and R. Mifflin), IIASA Proceedings 3, Pergamon, Oxford.

Lemarechal, C. and Mifflin, R. (eds) (1978). *Nonsmooth Optimization*, IIASA Proceedings 3, Pergamon, Oxford.

Lemke, C. E. (1965). Bimatrix equilibrium points and mathematical programming, *Management Sci.*, **11**, 681–689.

Levenberg, K. (1944). A method for the solution of certain nonlinear problems in least squares, *Quart. Appl. Math.*, **2**, 164–168.

Lustig, I. J. (1985). A practical approach to Karmarkar's algorithm, Stanford Univ. Systems Optimization Lab. Report SOL 85-5.

Madsen, K. (1975). An algorithm for minimax solution of overdetermined systems of nonlinear equations, *J. Inst. Maths. Applns.*, **16**, 321–328.

Maratos, N. (1978). Exact penalty function algorithms for finite dimensional and control optimization problems, Ph.D. thesis, Univ. of London.

Marquardt, D. W. (1963). An algorithm for least squares estimation of nonlinear parameters, *SIAM J.*, **11**, 431–441.

Marsten, R. E. (1975). The use of the boxstep method in discrete optimization, in *Nondifferentiable Optimization* (Eds M. L. Balinski and P. Wolfe), Mathematical Programming Study 3, North-Holland, Amsterdam.

Mayne, D. Q. (1980). On the use of exact penalty functions to determine step length in optimization algorithms, in *Numerical Analysis, Dundee 1979* (Ed. G. A. Watson), Lecture Notes in Mathematics 773, Springer-Verlag, Berlin.

McCormick, G. P. (1977). A modification of Armijo's step-size rule for negative curvature, *Math. Prog.*, **13**, 111–115.

McCormick. G. P. and Pearson, J. D. (1969). Variable metric methods and unconstrained optimization, in *Optimization* (Ed. R. Fletcher), Academic Press, London.

McLean, R. A. and Watson, G. A. (1980). Numerical methods for nonlinear discrete L_1 approximation problems, in *Numerical Methods of Approximation Theory* (Eds L. Collatz, G. Meinardus and H. Warner), ISNM 52, Birkhauser-Verlag, Basle.

Moré, J. J. (1978). The Levenberg–Marquardt algorithm: implementation and theory, in *Numerical Analysis*, Dundee 1977 (Ed. G. A. Watson), Lecture Notes in Mathematics 630, Springer-Verlag, Berlin.

Moré, J. J. and Sorensen, D. C. (1982). Newton's method, Argonne Nat. Lab. Report ANL-82-8.

Morrison, D. D. (1968). Optimization by least squares, *SIAM J. Num. Anal.*, **5**, 83–88.

Murray, W. (1969). An algorithm for constrained minimization, in *Optimization* (Ed. R. Fletcher), Academic Press, London.

Murray, W. (1971). An algorithm for finding a local minimum of an indefinite quadratic program, NPL Report NAC 1.

Murray, W. (1972). Second derivative methods, in *Numerical Methods for Unconstrained Optimization* (Ed. W. Murray), Academic Press, London.

Murtagh, B. A. and Sargent, R. W. H. (1969). A constrained minimization method with quadratic convergence, in *Optimization* (Ed. R. Fletcher), Academic Press, London.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization, *Computer J.*, **7**, 308–313.

Nocedal, J. and Overton, M. L. (1985). Projected Hessian updating algorithms for nonlinearly constrained optimization, *SIAM J. Num. Anal.*, **22**, 821–850.

Oren, S. S. (1974). On the selection of parameters in self-scaling variable metric algorithms, *Math. Prog.*, **7**, 351–367.

Osborne, M. R. (1972). Topics in optimization, Stanford Univ. Dept. of Comp. Sci. Report STAN-CS-72-279.

Osborne, M. R. (1985). *Finite Algorithms in Optimization and Data Analysis*, John Wiley, Chichester.

Osborne, M. R. and Watson, G. A. (1969). An algorithm for minimax approximation in the nonlinear case, *Computer J.*, **12**, 63–68.

Pietrzykowski, T. (1969). An exact potential method for constrained maxima, *SIAM J. Num. Anal.*, **6**, 217–238.

Polak, E. (1971). *Computational Methods in Optimization: A Unified Approach*, Academic Press, New York.

Powell, M. J. D. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, **7**, 155–162.

Powell, M. J. D. (1965). A method for minimizing a sum of squares of nonlinear functions without calculating derivatives, *Computer J.*, **11**, 302–304.

Powell, M. J. D. (1969). A method for nonlinear constraints in minimization problems, in *Optimization* (Ed. R. Fletcher), Academic Press, London.

Powell, M. J. D. (1970a). A new algorithm for unconstrained optimization, in *Nonlinear Programming* (Eds J. B. Rosen, O. L. Mangasarian and K. Ritter), Academic Press, New York.

Powell, M. J. D. (1970b). A hybrid method for nonlinear equations, in *Numerical Methods for Nonlinear Algebraic Equations* (Ed. P. Rabinowitz), Gordon and Breach, London.

Powell, M. J. D. (1971). On the convergence of the variable metric algorithm, *J. Inst. Maths. Applns.*, **7**, 21–36.

Powell, M. J. D. (1972a). Unconstrained minimization and extensions for constraints, AERE Harwell Report TP495.

Powell, M. J. D. (1972b). Some properties of the variable metric algorithm, in *Numerical Methods for Nonlinear Optimization* (Ed. F. A. Lootsma), Academic Press, London.

Powell, M. J. D. (1972c). Quadratic termination properties of minimization algorithms, I and II, *J. Inst. Maths. Applns.*, **10**, 333–342 and 343–357.

Powell, M. J. D. (1972d). Unconstrained minimization algorithms without computation of derivatives, AERE Harwell Report TP483.

Powell, M. J. D. (1972e). Problems related to unconstrained optimization, in *Numerical Methods for Unconstrained Optimization* (Ed. W. Murray), Academic Press, London.

Powell, M. J. D. (1973). On search directions for minimization algorithms, *Math. Prog.*, **4**, 193–201.

Powell, M. J. D. (1975a). A view of minimization algorithms that do not require derivatives, *ACM Trans. Math. Software*, **1**, 97–107.

Powell, M. J. D. (1975b). Convergence properties of a class of minimization algorithms, in *Nonlinear Programming 2* (Eds O. L. Mangasarian, R. R. Meyer and S. M. Robinson, Academic Press, New York.

Powell, M. J. D. (1976). Some global convergence properties of a variable metric algorithm for minimization without exact line searches, in *SIAM-AMS Proceedings, Vol. IX* (Eds R. W. Cottle and C. E. Lemke), SIAM Publications, Philadelphia.

Powell, M. J. D. (1977a). Quadratic termination properties of Davidon's new variable metric algorithm, *Math. Prog.*, **12**, 141–147.

Powell, M. J. D. (1977b). Restart procedures for the conjugate gradient method, *Math. Prog.*, **12**, 241–254.

Powell, M. J. D. (1977c). Constrained optimization by a variable metric method, Cambridge Univ. DAMTP Report 77/NA6.

Powell, M. J. D. (1978a). A fast algorithm for nonlinearly constrained optimization calculations, in *Numerical Analysis, Dundee 1977* (Ed. G. A. Watson), Lecture Notes in Mathematics 630, Springer-Verlag, Berlin.

Powell, M. J. D. (1978b). The convergence of variable metric methods for nonlinearly constrained optimization calculations, in *Nonlinear Programming 3* (Eds O. L. Mangasarian, R. R. Meyer and S. M. Robinson), Academic Press, New York.

Powell, M. J. D. (1985a). On the quadratic programming algorithm of Goldfarb and Idnani, in *Mathematical Programming Essays in Honor of George B. Dantzig, Part II* (Ed. R. W. Cottle), Mathematical Programming Study 25, North-Holland, Amsterdam.

Powell, M. J. D. (1985b). On error growth in the Bartels–Golub and Fletcher–Matthews algorithms for updating matrix factorizations, Cambridge Univ. DAMTP Report 85/NA10.

Powell, M. J. D. (1985c). The performance of two subroutines for constrained optimization on some difficult test problems, in *Numerical Optimization 1984* (Eds P. T. Boggs, R. H. Byrd and R. B. Schnabel), SIAM Publications, Philadelphia.

Powell, M. J. D. (1985d). How bad are the BFGS and DFP methods when the objective function is quadratic?, Univ. of Cambridge DAMTP Report 85/NA4.

Powell, M. J. D. (1985e). Updating conjugate directions by the BFGS formula, Univ. of Cambridge DAMTP Report 85/NA11.

Powell, M. J. D. and Yuan, Y. (1986). A recursive quadratic programming algorithm that uses differentiable penalty functions, *Math. Prog.*, **35**, 265–278.

Pshenichnyi, B. N. (1978). Nonsmooth optimization and nonlinear programming, in *Nonsmooth Optimization* (Eds C. Lemarechal and R. Mifflin), IIASA Proceedings 3, Pergamon, Oxford.

Rall, L. B. (1969). *Computational Solution of Nonlinear Operator Equations*, John Wiley, New York.

Reid, J. K. (1971a). On the method of conjugate gradients for the solution of large sparse systems of equations, in *Large Sparse Sets of Linear Equations* (Ed. J. K. Reid), Academic Press, London.

Reid, J. K. (ed.) (1971b). *Large Sparse Sets of Linear Equations*, Academic Press, London.

Reid, J. K. (1975). A sparsity-exploiting version of the Bartels–Golub decomposition for linear programming bases, AERE Harwell Report CSS20.

Rhead, D. (1974). On a new class of algorithms for function minimization without evaluating derivatives, Univ. of Nottingham, Dept. of Mathematics Report.

Robinson, S. M. (1982). Generalised equations and their solutions, Part II: applications to nonlinear programming, in *Optimality and Stability in Mathematical Programming* (Ed. M. Guignard), Mathematical Programming Study 19, North-Holland, Amsterdam.

Rockafeller, R. T. (1974). Augmented Lagrange multiplier functions and duality in non-convex programming, *SIAM J. Control*, **12**, 268–285.

Rosen, J. B. (1960). The gradient projection method for nonlinear programming, Part I: Linear constraints, *J. SIAM*, **8**, 181–217.

Rosen, J. B. (1961). The gradient projection method for nonlinear programming, Part II: Nonlinear constraints, *J. SIAM*, **9**, 514–532.

Sargent, R. W. H. (1974). Reduced gradient and projection methods for nonlinear programming, in *Numerical Methods for Constrained Optimization* (Eds P. E. Gill and W. Murray), Academic Press, London.

Saunders, M. A. (1972). Product form of the Cholesky factorization for large-scale linear Programming, Stanford Univ. Report STAN-CS-72-301.

Schubert, L. K. (1970). Modification of a quasi-Newton method for nonlinear equations with a sparse Jacobian, *Maths. Comp.*, **24**, 27–30.

Schittkowski, K. (1983). On the convergence of a sequential quadratic programming method with an augmented Lagrangian line search function, *Math. Operationsforschung u. Stat., Ser. Optimization*, **14**, 197–216.

Schittkowski, K. and Stoer, J. (1979). A factorization method for the solution of constrained linear least-squares problems allowing data changes, *Numer. Math.*, **31**, 431–463.

Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization, *Maths. Comp.*, **24**, 647–656.

Shanno, D. F. and Phua, K.-H. (1978). Matrix conditioning and nonlinear optimization, *Math. Prog.*, **14**, 149–160.

Sherman, A. H. (1978). On Newton iterative methods for the solution of systems of nonlinear equations, *SIAM J. Num. Anal.*, **15**, 755–771.

Sinclair, J. E. and Fletcher, R. (1974). A new method of saddle point location for the calculation of defect migration energies, *J. Phys. C: Solid State Phys.*, **7**, 864–870.

Sinclair, J. W. (1979). On quasi-Newton methods, Ph.D. thesis, Univ. of Dundee Dept. of Mathematics.

Smith, C. S. (1962). The automatic computation of maximum likelihood estimates, N.C.B. Sci. Dept. Report SC846/MR/40.

Spendley, W., Hext, G. R. and Himsworth, F. R. (1962). Sequential application of simplex designs in optimization and evolutionary operation, *Technometrics*, **4**, 441–461.

Stewart, G. W. (1967). A modification of Davidon's minimization method to accept difference approximations to derivatives, *J. Ass. Comput. Mach.*, **14**, 72–83.

Stoer, J. (1975). On the convergence rate of imperfect minimization algorithms in Broyden's β-class, *Math. Prog.*, **9**, 313–335.

Swann, W. H. (1972). Direct search methods, in *Numerical Methods for Unconstrained Optimization* (Ed. W. Murray), Academic Press, London.

Swann, W. H. (1974). Constrained optimization by direct search, in *Numerical Methods for Constrained Optimization* (Eds P. E. Gill and W. Murray), Academic Press, London.

Tapia, R. (1984). On the characterization of Q-superlinear convergence of quasi-Newton methods for constrained optimization, Rice Univ. Dept. of Math. Sci. Report 84–2.

Toint, Ph.L. (1977). On sparse and symmetric updating subject to a linear equation, *Maths. Comp.*, **31**, 954–961.

Varga, R. S. (1962). *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J.

Walsh, G. R. (1975). *Methods of Optimization*, John Wiley, London.

Watson, G. A. (1978). A class of programming problems whose objective function contains a norm, *J. Approx. Theo.*, **23**, 401–411.

Watson, G. A. (1979). The minimax solution of an overdetermined system of nonlinear equations, *J. Inst. Maths. Applns.*, **23**, 167–180.

Wilkinson, J. H. (1965). *The Algebraic Eigenvalue Problem*, Oxford Univ. Press, Oxford.

Wilkinson, J. H. (1977). Some recent advances in numerical linear algebra, in *The State of the Art in Numerical Analysis* (Ed. D. A. H. Jacobs), Academic Press, London.

Wilson, R. B. (1963). A simplicial algorithm for concave programming, Ph.D. dissertation, Harvard Univ. Graduate School of Business Administration.

Winfield, D. (1973). Function minimization by interpolation in a data table, *J. Inst. Maths. Applns.*, **12**, 339–348.

Wolfe, P. (1959). The secant method for simultaneous nonlinear equations, *Comm. Ass. Comput. Mach.*, **2**, 12–13.

Wolfe, P. (1961). A duality theorem for nonlinear programming, *Quart. Appl. Math.*, **19**, 239–244.

Wolfe, P. (1963a). Methods of nonlinear programming, in *Recent Advances in Mathematical Programming* (Eds R. L. Graves and P. Wolfe), McGraw-Hill, New York.

Wolfe, P. (1963b). A technique for resolving degeneracy in linear programming, *J. SIAM*, **11**, 205–211.

Wolfe, P. (1965). The composite simplex algorithm, *SIAM Rev.*, **7**, 42–54.

Wolfe, P. (1967). Methods of nonlinear programming, in *Nonlinear Programming* (Ed. J. Abadie), North-Holland, Amsterdam.

Wolfe, P. (1968a). Another variable metric method, working paper.

Wolfe, P. (1968b). Convergence conditions for ascent methods, *SIAM Rev.*, **11**, 226–235.

Wolfe, P. (1972). On the convergence of gradient methods under constraint, *IBM J. Res. and Dev*, **16**, 407–411.

Wolfe, P. (1975). A method of conjugate subgradients, in *Nondifferentiable Optimization* (Eds M. L. Balinski and P. Wolfe), Mathematical Programming Study 3, North-Holland, Amsterdam.

Wolfe, P. (1980). The ellipsoid algorithm, *Optima* (Math. Prog. Soc. newsletter), Number 1.

Womersley, R. S. (1978). An approach to nondifferentiable optimization, M.Sc. thesis, Univ. of Dundee Dept. of Mathematics.

Womersley, R. S. (1981). Numerical methods for structured problems in nonsmooth optimization, Ph.D. thesis, Univ. of Dundee Dept. of Mathematics.

Womersley, R. S. (1982). Optimality conditions for piecewise smooth functions, in

Nondifferential and Variational Techniques in Optimization (Eds D. C. Sorensen and R. J.-B. Wets), Mathematical Programming Study 17, North-Holland, Amsterdam.

Womersley, R. S. (1984a). Censored discrete linear l_1 approximation, ANU Dept. of Math. Sci. Res. Report No. 8.

Womersley, R. S. (1984b). Minimizing nonsmooth composite functions, ANU Dept. of Math. Sci. Res. Report No. 13.

Womersley, R. S. and Fletcher, R. (1986). An algorithm for composite nonsmooth optimization problems, *J. Opt. Theo. Applns.*, **48**, 493–523.

Yuan, Y. (1984). An example of only linear convergence of trust region algorithms for nonsmooth optimization, *IMA J. Num. Anal.*, **4**, 327–335.

Yuan, Y. (1985a). Conditions for convergence of trust region algorithms for nonsmooth optimization, *Math. Prog.*, **31**, 220–228.

Yuan, Y. (1985b). On the superlinear convergence of a trust region algorithm for nonsmooth optimization, *Math. Prog.*, **31**, 269–285.

Zoutendijk, G. (1960). *Methods of Feasible Directions*, Elsevier, Amsterdam.

Chapter 14

Non-Smooth Optimization

14.1 INTRODUCTION

In Part 1 on unconstrained optimization an early assumption is to exclude all problems for which the objective function $f(\mathbf{x})$ is not smooth. Yet practical problems sometimes arise (*non-smooth optimization* (NSO) or *non-differentiable optimization* problems) which do not meet this requirement; this chapter studies progress which has been made in the practical solution of such problems. Examples of NSO problems occur when solving nonlinear equations $c_i(\mathbf{x}) = 0$, $i = 1, 2, \ldots, m$ (see Section 6.1) by minimizing $\|\mathbf{c}(\mathbf{x})\|_1$ or $\|\mathbf{c}(\mathbf{x})\|_\infty$. This arises either when solving simultaneous equations exactly ($m = n$) or when finding best solutions to over-determined systems ($m > n$) as in data fitting applications. Another similar problem is that of finding a feasible point of a system of nonlinear inequalities $c_i(\mathbf{x}) \leqslant 0$, $i = 1, 2, \ldots, m$, by minimizing $\|\mathbf{c}(\mathbf{x})^+\|_1$ or $\|\mathbf{c}(\mathbf{x})^+\|_\infty$ where $c_i^+ = \max(c_i, 0)$. A generalization of these problems arises when the equations or inequalities are the constraints in a nonlinear programming problem (Chapter 12). Then a possible approach is to use an exact penalty function and minimize functions like $\nu f(\mathbf{x}) + \|\mathbf{c}(\mathbf{x})\|$ or $\nu f(\mathbf{x}) + \|\mathbf{c}(\mathbf{x})^+\|$, in particular using the L_1 norm. This idea is attracting much research interest and is expanded upon in Section 14.3. Yet another type of problem is to minimize the *max function* $\max_i c_i(\mathbf{x})$ where the max is taken over some finite set. This includes many examples from electrical engineering including microwave network design and digital filter design (Charalambous, 1979). In fact almost all these examples can be considered in a more generalized sort of way as special cases of a certain *composite function* and a major portion of this chapter is devoted to studying this type of function. In particular the term *composite NSO* is used to describe this type of problem.

Another common source of NSO problems arises when using the decomposition principle. For example the LP problem

$$\begin{array}{ll} \underset{\mathbf{x},\mathbf{y}}{\text{minimize}} & \mathbf{c}^{\mathrm{T}}\mathbf{x} + \mathbf{d}^{\mathrm{T}}\mathbf{y} \\ \text{subject to} & \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y} \leqslant \mathbf{b} \end{array}$$

can be written as the convex NSO problem

$$\underset{\mathbf{x}}{\text{minimize}}\, f(\mathbf{x}) \triangleq \mathbf{c}^{\mathrm{T}}\mathbf{x} + \min_{\mathbf{y}} [\mathbf{d}^{\mathrm{T}}\mathbf{y} : \mathbf{B}\mathbf{y} \leqslant \mathbf{b} - \mathbf{A}\mathbf{x}].$$

Application of NSO methods to column generation problems and to a variety of scheduling problems is described by Marsten (1975) and the application to network scheduling problems is described by Fisher, Northup, and Shapiro (1975). Both these papers appear in the Mathematical Programming Study 3 (Balinski and Wolfe, 1975) which is a valuable reference. Another good source is the book by Lemarechal and Mifflin (1978) and their test problems 3 and 4 illustrate further applications.

Insofar as I understand them, all these applications could in principle be formulated as max functions and hence as composite functions (see below). However in practice this would be too complicated or require too much storage. Thus a different situation can arise in which the only information available at any point $\mathbf{x}$ is $f(\mathbf{x})$ and a vector $\mathbf{g}$. Usually $\mathbf{g}$ is ∇f, but if f is non-differentiable at $\mathbf{x}$ then $\mathbf{g}$ is an element of the subdifferential ∂f (see Section 14.2) for convex problems, or the generalized gradient in non-convex problems. Since less information about $f(\mathbf{x})$ is available, this type of application is more difficult than the composite function application and is referred to here as *basic NSO*. However, both types of application have some structure in common.

In view of this common structure, algorithms for all types of NSO problem are discussed together in Section 14.4. For composite NSO it is shown that an algorithm with second order convergence can be obtained by making linearizations of the individual functions over which the composition is taken, together with an additional (smooth) quadratic approximation which includes curvature information from the functions in the composition. Variants of this algorithm have been used in a number of practical applications. A potential difficulty for this algorithm (and indeed for any NSO algorithm), known as the *Maratos effect* is described, in which the expected superlinear convergence may not be obtained. It is shown how this arises, and that it can be circumvented in a simple way by adding a *second order correction step* to the algorithm. This algorithm can also be made globally convergent by incorporating the idea of a trust region and details of this result are given in Section 14.5. The algorithm is applicable to both max function and L_1 and L_∞ approximation problems and to nonlinear programming applications of non-smooth exact penalty functions. Algorithms for basic NSO have not progressed as far because of the difficulties caused by the limited availability of information. One type of method—the *bundle method*—tries to accumulate information locally about the subdifferential of the objective function. Alternatively linearization methods have also been tried and the possibility of introducing second order information into these algorithms is being considered. In fact there is currently much research interest in NSO algorithms of all kinds and further developments can be expected.

A prerequisite for describing NSO problems and algorithms is a study of optimality conditions for non-smooth functions. This can be done at various levels of generality. I have tried to make the approach here as simple and readable as possible (although this is not easy in view of the inherent difficulty of the material), whilst trying to cover all practical applications. The chapter is largely self-contained and does not rely on any key results from outside sources. One requirement is the extension of the material in Section 9.4 on convex functions to include the non-smooth case. The concept of a subdifferential is introduced and the resulting definitions of directional derivatives lead to a simple statement of first order conditions. More general applications can be represented by the *composite function*

$$\phi(\mathbf{x}) \triangleq f(\mathbf{x}) + h(\mathbf{c}(\mathbf{x})) \tag{14.1.1}$$

and the problem of minimizing such a function is referred to as *composite NSO*. Here $f(\mathbf{x})$ ($\mathbb{R}^n \to \mathbb{R}^1$) and $\mathbf{c}(\mathbf{x})$ ($\mathbb{R}^n \to \mathbb{R}^m$) are smooth functions ($\mathbb{C}^1$), and $h(\mathbf{c})$ ($\mathbb{R}^m \to \mathbb{R}^1$) is convex but non-smooth ($\mathbb{C}^0$). Note that it would be possible more simply to express $\phi(\mathbf{x}) = h(\mathbf{c}(\mathbf{x}))$ and to regard (14.1.1) merely as a special case of this. I have not done this because in penalty function applications it is much more convenient to use (14.1.1). Other applications in which $f(\mathbf{x})$ is not present are readily handled by setting $f(\mathbf{x}) = 0$. Important special cases of $h(\mathbf{c})$ are set out in (14.1.3) below. Optimality conditions for (14.1.1) can be obtained as a straightforward extension of those for convex functions. This material is set out in Section 14.2 and both first and second order conditions are studied. An important aspect of this theory is to give a regularity condition under which the second order conditions are almost necessary and sufficient. This involves the idea of *structure functionals*. It is shown that these functionals enable an equivalence to be drawn in the neighbourhood of a local solution between a composite NSO problem and a nonlinear programming problem whose constraints involve the structure functionals.

There are NSO problems which do not fit into the category of composite NSO (for example $\phi(\mathbf{x}) = \max(0, \min(x_1, x_2))$), and which are not even well modelled locally by (14.1.1). A practical example is that of *censored L_1 approximation* (Womersley, 1984a), but few examples of this type occur in practice. Wider classes of function have been introduced which cover these cases, for example the *locally Lipschitz functions* of Clarke (1975), but substantial complication in the analysis is introduced. Also these classes do not directly suggest algorithms as the function (14.1.1) does. Furthermore the resulting first order conditions permit descent directions to occur and so are of little practical use (Womersley, 1982). Hence these classes are not studied any further here. Finally the problem of *constrained NSO* can be considered. First order conditions for smooth constraint functions are given by Watson (1978) and for constrained L_∞ approximation by Andreassen and Watson (1976). First and second order conditions for a single non-smooth constraint function involving a norm are given by Fletcher and Watson (1980). In Section 14.6 these results are extended to cover a composite objective function and a single composite constraint

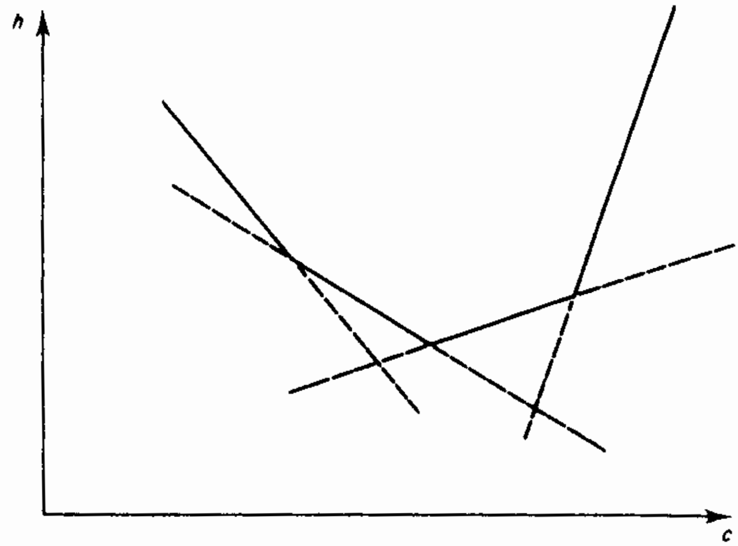

Figure 14.1.1 Polyhedral convex function

function (*constrained composite NSO*), and this is the most general case that needs to be considered. A regularity condition involving the use of structure functionals is also given.

In most cases (but not all—for example the exact penalty function $\phi = \nu f + \|\mathbf{c}\|_2$) a further specialization can be made in which $h(\mathbf{c})$ is restricted to be a *polyhedral convex function.* In this case the graph of $h(\mathbf{c})$ is made up of a finite number of supporting hyperplanes $\mathbf{c}^T\mathbf{h}_i + b_i$, and $h(\mathbf{c})$ is thus defined by

$$h(\mathbf{c}) \triangleq \max_i (\mathbf{c}^T\mathbf{h}_i + b_i) \tag{14.1.2}$$

where the vectors $\mathbf{h}_i$ and scalars b_i are given. Thus the polyhedral convex function is a max function in terms of linear combinations of the elements of $\mathbf{c}$. A simple illustration is given in Figure 14.1.1. Most interest lies in five special cases of (14.1.2), in all of which $b_i = 0$ for all i. These functions are set out below, together with the vectors $\mathbf{h}_i$ which define them given as columns of a matrix $\mathbf{H}$.

Case I	$h(\mathbf{c}) = \max_i c_i$	$\mathbf{H} = \mathbf{I}$ $\quad (m \times m)$	
Case II	$h(\mathbf{c}) = \|\mathbf{c}^+\|_\infty$	$\mathbf{H} = [\mathbf{I}:\mathbf{0}]$ $\quad (m \times (m+1))$	
Case III	$h(\mathbf{c}) = \|\mathbf{c}\|_\infty$	$\mathbf{H} = [\mathbf{I}: -\mathbf{I}]$ $\quad (m \times 2m)$	
Case IV	$h(\mathbf{c}) = \|\mathbf{c}^+\|_1$	columns of $\mathbf{H}$ are all possible combinations of 1 and 0 $\quad (m \times 2^m)$	(14.1.3)
Case V	$h(\mathbf{c}) = \|\mathbf{c}\|_1$	columns of $\mathbf{H}$ are all possible combinations of 1 and -1 $\quad (m \times 2^m)$	

For example with $m = 2$ in case (V) the matrix $\mathbf{H}$ is $\mathbf{H} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$.
Contours of these functions for $m = 2$ are illustrated in Figure 14.1.2. The broken lines indicate the surfaces (*grooves*) along which the different linear pieces join up and along which the derivative is discontinuous. The term *piece* is used to

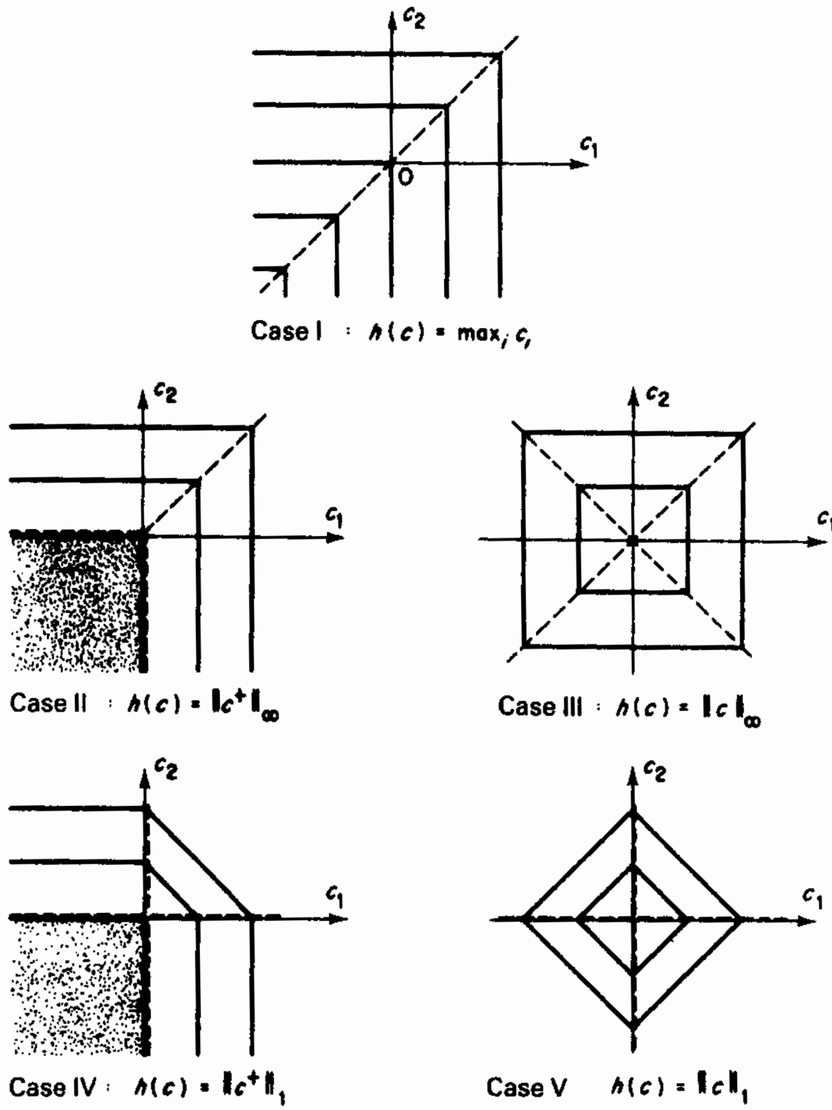

Figure 14.1.2 Contours of polyhedral convex functions

denote that part of the graph of a max function in which one particular function achieves the maximum value, for example the graph illustrated in case (II) of Figure 14.1.2 is made up of three linear pieces and the active equations are $h(\mathbf{c}) = c_2$ for $c_1 \leqslant c_2$ and $c_2 \geqslant 0$, $h(\mathbf{c}) = c_1$ for $c_2 \leqslant c_1$ and $c_1 \geqslant 0$, and $h(c) = 0$ for $c_1 \leqslant 0$ and $c_2 \leqslant 0$. The subdifferential $\partial h(\mathbf{c})$ for a polyhedral convex function has the simple form expressed in Lemma 14.2.2 as the convex hull of the gradients of the active pieces. The polyhedral nature of $h(\mathbf{c})$ has important consequences in regard to second order conditions, and this is considered later in Section 14.2.

In fact it is possible to derive optimality results for the composite function $\phi(\mathbf{x})$ in (14.1.1) when $h(\mathbf{c})$ is the polyhedral function (14.1.2) without using notions of convexity at all. Clearly $\phi(\mathbf{x})$ can be written

$$\phi(\mathbf{x}) = \max_i \left(f(\mathbf{x}) + \mathbf{c}(\mathbf{x})^{\mathrm{T}}\mathbf{h}_i + b_i\right) \tag{14.1.4}$$

which is equivalent to

$$\phi(\mathbf{x}) = \min v\colon v \geqslant f(\mathbf{x}) + \mathbf{c}(\mathbf{x})^{\mathrm{T}}\mathbf{h}_i + b_i \qquad \forall i. \tag{14.1.5}$$

Therefore $\mathbf{x}^*$ minimizes $\phi(\mathbf{x})$ locally iff $\mathbf{x}^*, v^*$ is a local solution of the nonlinear programming problem

$$\begin{aligned} &\underset{\mathbf{x},v}{\text{minimize}} \quad v \\ &\text{subject to} \quad v - f(\mathbf{x}) - \mathbf{c}(\mathbf{x})^{\mathrm{T}}\mathbf{h}_i \geqslant b_i \qquad \forall i. \end{aligned} \tag{14.1.6}$$

Thus first and second order optimality conditions for (14.1.4) can be obtained by applying the equivalent results in nonlinear programming to (14.1.6), that is Theorem 9.1.1, 9.3.1, and 9.3.2. It turns out however that this approach is less general and somewhat clumsy, so it is only sketched out briefly here, and the derivation given in Section 14.2 is preferred. The first difficulty concerns the regularity assumption $\mathscr{F}' = F'$. It is possible (but not trivial—see Question 14.1) to prove that this always holds, and therefore it is important not to make an unnecessary independence assumption. The appropriate Lagrangian function for Theorem 9.1.1 is

$$\mathscr{L}(\mathbf{x}, v, \boldsymbol{\mu}) = v - \sum_i \mu_i (v - f(\mathbf{x}) - \mathbf{c}(\mathbf{x})^{\mathrm{T}}\mathbf{h}_i - b_i) \tag{14.1.7}$$

and it is implied at $\mathbf{x}^*, v^*$ that multipliers $\boldsymbol{\mu}^*$ exist such that

$$\begin{aligned} &\frac{\partial}{\partial v}\mathscr{L}(\mathbf{x}^*, v^*, \boldsymbol{\mu}^*) = 0 \quad \text{or} \quad \sum \mu_i^* = 1 \\ &\nabla_x \mathscr{L}(\mathbf{x}^*, v^*, \boldsymbol{\mu}^*) = \mathbf{0} \quad \text{or} \quad \sum \mu_i^*(\mathbf{g}^* + \mathbf{A}^*\mathbf{h}_i) = \mathbf{0} \\ &\boldsymbol{\mu}^* \geqslant \mathbf{0} \\ &\mu_i^* > 0 \Rightarrow v^* = f^* + \mathbf{c}^{*\mathrm{T}}\mathbf{h}_i + b_i. \end{aligned} \tag{14.1.8}$$

As in previous chapters the notation $\mathbf{g} = \nabla f$ and $\mathbf{A} = \nabla \mathbf{c}^{\mathrm{T}}$ is used and $f^* = f(\mathbf{x}^*)$, etc. If $\boldsymbol{\lambda} = \mathbf{H}\boldsymbol{\mu}$ is written, then the existence of a vector $\boldsymbol{\mu}^*$ in the above is equivalent to the existence of a vector $\boldsymbol{\lambda}^*$ in the set

$$\partial h^* = \underset{i \in \mathscr{A}^*}{\operatorname{conv}} \mathbf{h}_i \tag{14.1.9}$$

such that $\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}$. In fact ∂h^* is the subdifferential of $h(\mathbf{c}^*)$ as shown in Lemma 14.2.2. $\mathscr{A}^*$ is the set of active constraints or equivalently the set of indices at which the max in (14.1.4) is attained, that is

$$\mathscr{A}^* = \{i\colon \mathbf{c}^{*\mathrm{T}}\mathbf{h}_i + b_i = h(\mathbf{c}^*)\}. \tag{14.1.10}$$

It is possible to use (14.1.9) to obtain equivalent but more convenient expressions for $\partial h(\mathbf{c})$ in cases (I) to (V) in (14.1.3) above as follows:

$$\partial \max_i c_i = \{\boldsymbol{\lambda}\colon \sum \lambda_i = 1,\ \boldsymbol{\lambda} \geqslant \mathbf{0},\ c_i < \max_i c_i \Rightarrow \lambda_i = 0\} \tag{14.1.11}$$

$$\begin{aligned} \partial \|\mathbf{c}^+\|_\infty = \{\boldsymbol{\lambda}\colon \sum \lambda_i = 1,\ \boldsymbol{\lambda} \geqslant \mathbf{0},\ c_i < \|\mathbf{c}^+\|_\infty \Rightarrow \lambda_i = 0 \\ \mathbf{c}^+ \neq \mathbf{0} \Rightarrow \sum \lambda_i = 1\} \end{aligned} \tag{14.1.12}$$

$$\partial\|\mathbf{c}\|_\infty = \{\boldsymbol{\lambda}\colon \mathbf{c}\neq\mathbf{0} \Rightarrow \sum|\lambda_i| = 1$$
$$|c_i| < \|\mathbf{c}\|_\infty \Rightarrow \lambda_i = 0$$
$$|c_i| = \|\mathbf{c}\|_\infty \Rightarrow \lambda_i c_i \geqslant 0$$
$$\mathbf{c}=\mathbf{0} \Rightarrow \sum|\lambda_i| \leqslant 1\} \tag{14.1.13}$$

$$\partial\|\mathbf{c}^+\|_1 = \{\boldsymbol{\lambda}\colon 0\leqslant\lambda_i\leqslant 1,\ c_i>0 \Rightarrow \lambda_i = 1$$
$$c_i<0 \Rightarrow \lambda_i = 0\} \tag{14.1.14}$$

$$\partial\|\mathbf{c}\|_1 = \{\boldsymbol{\lambda}\colon |\lambda_i|\leqslant 1,\ c_i\neq 0 \Rightarrow \lambda_i = \operatorname{sign} c_i\}. \tag{14.1.15}$$

The equivalence between these sets and those defined by $\operatorname{conv}_{i\in\mathscr{A}}\mathbf{h}_i$ is left as an exercise (Question 14.3). Expressions (14.1.12) to (14.1.15) can also be derived as special cases of (14.3.7) or (14.3.8).

It is also possible to apply the development of (9.3.9) onwards to problem (14.1.6) to obtain the second order conditions of Theorem 14.2.2 and 14.2.3, and also the regularity condition (14.2.31) (see Lemma 14.2.7). To do this an index $p\in\mathscr{A}_+^*$ is used to eliminate v (or s_{n+1}—see Question 14.1). The analysis is not entirely straightforward so again the more direct and more general approach of Section 14.2 is preferred.

The parameters $\boldsymbol{\lambda}^*\in\partial h^*$ which exist at the solution to an NSO problem are closely related to Lagrange multipliers and indeed they can be given a simple interpretation similar to (9.1.10). Consider a perturbed problem in which $\mathbf{c}(\mathbf{x})$ is replaced by $\mathbf{c}(\mathbf{x})+\boldsymbol{\varepsilon}$ giving a function $\phi_\varepsilon(\mathbf{x})$ and assume that the minimizer $\mathbf{x}(\boldsymbol{\varepsilon})$ is such that the same constraints are active in (14.1.6). It follows that the right-hand side of each active constraint is perturbed by an amount $\boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{h}_i$. Since μ_i^* is the multiplier of the ith constraint and using (9.1.10) it follows that the change to v is $\Delta v = \sum_i \boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{h}_i\mu_i^* = \boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\lambda}^*$ and hence that

$$\frac{\mathrm{d}\phi^*}{\mathrm{d}\varepsilon_i} = \lambda_i^*. \tag{14.1.16}$$

Thus λ_i^* measures the first order rate of change of the optimum function value consequent on perturbations to c_i. This result can also be used to illustrate the need for the condition $\boldsymbol{\lambda}\geqslant\mathbf{0}$ in (14.1.11). For example if the first order conditions arising from (14.1.8) ($\exists\boldsymbol{\lambda}^*\in\partial h^*$ such that $\mathbf{g}^*+\mathbf{A}^*\boldsymbol{\lambda}^*=\mathbf{0}$) are satisfied except that $\lambda_i^*<0$ for some i, then it is possible to show that $\mathbf{x}^*$ is not optimal. Consider a perturbation $\boldsymbol{\varepsilon}$ with $\varepsilon_i>0$ and $\varepsilon_j = 0, j\neq i$. Existence of the solution $\mathbf{x}(\boldsymbol{\varepsilon})$ follows under a suitable independence assumption using the implicit function theorem. At $\mathbf{x}(\boldsymbol{\varepsilon})$ the max in (14.1.2) is achieved by all $j\in\mathscr{A}^*$ if $\boldsymbol{\varepsilon}$ is small enough, so for all $j\in\mathscr{A}^*, j\neq i$,

$$c_j(\mathbf{x}(\boldsymbol{\varepsilon})) = c_i(\mathbf{x}(\boldsymbol{\varepsilon})) + \varepsilon_i.$$

Thus $c_i(\mathbf{x}(\boldsymbol{\varepsilon})) < c_j(\mathbf{x}(\boldsymbol{\varepsilon}))$ and hence $\phi_\varepsilon(\mathbf{x}(\boldsymbol{\varepsilon})) = \phi_0(\mathbf{x}(\boldsymbol{\varepsilon}))$ where ϕ_0 refers to the unperturbed function. It follows from (14.1.16) that $\mathrm{d}\phi_0^*/\mathrm{d}\varepsilon_i = \lambda_i^* < 0$ and so $\mathbf{x}^*$ is not a local solution. The situation is illustrated in Figure 14.1.3. Clearly increasing c_1 by ε does not reduced $\phi(\mathbf{x})$ in case (i) when $\boldsymbol{\lambda}\geqslant\mathbf{0}$ but does reduce $\phi(\mathbf{x})$ in case (ii). Another example of this is given in Question 14.12.

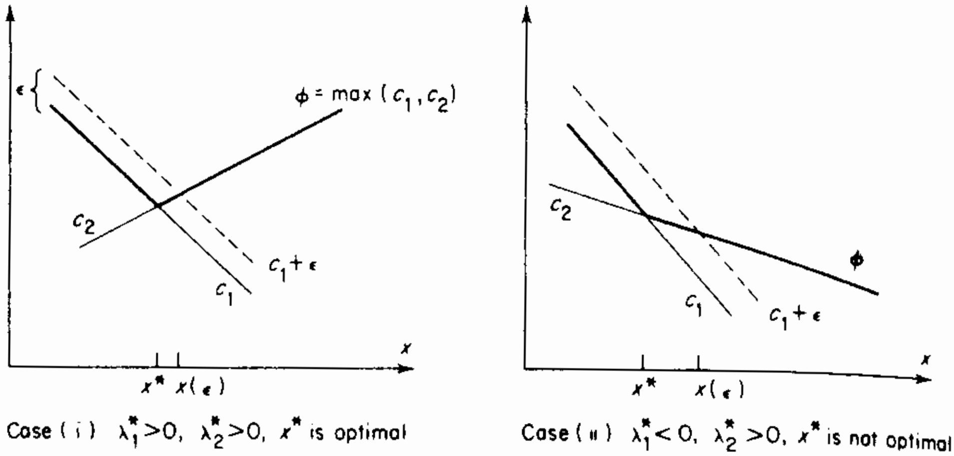

Figure 14.1.3. Interpretation of multipliers in NSO

The requirement that the active constraints in (14.1.6) remain the same is important and is in the nature of an independence assumption. This usually holds for case (I) and (II) problems in (14.1.3) and for case (III) problems when $\mathbf{c}^* \neq \mathbf{0}$. However for the L_1 norm functions of cases (IV) and (V) the assumption is not likely to hold. An alternative interpretation of Lagrange multipliers suitable for the L_1 case is given in Section 12.3. In Section 14.2 it is shown how all these cases, and also some others not involving polyhedral $h(\mathbf{c})$, can be handled within a unified framework by using a regularity assumption based on the concept of structure functionals.

14.2 OPTIMALITY CONDITIONS

In this section a number of results are proved leading to optimality conditions for the composite functions described in the previous section. Since these functions are defined in terms of non-smooth convex functions it is important to study the latter case first. Such a function is illustrated in Figure 14.2.1 and it can be seen that the non-differentiability at $\mathbf{x}^*$ allows the possibility of a number of supporting hyperplanes, in contrast to Figure 9.4.2. (Supporting hyperplanes to the epigraph of a convex function, which is a convex set, always exist; see Hadley, 1961.) Given a convex function $f(\mathbf{x})$ defined on a convex set $K \subset \mathbb{R}^n$, and given $\mathbf{x} \in \text{interior}\ (K)$, then for each supporting hyperplane at $\mathbf{x}$ it follows that an inequality of the form

$$f(\mathbf{x}+\boldsymbol{\delta}) \geqslant f(\mathbf{x}) + \boldsymbol{\delta}^{\mathrm{T}}\mathbf{g} \qquad \forall\, \mathbf{x}+\boldsymbol{\delta} \in K \tag{14.2.1}$$

holds, and $\mathbf{g}$ is a normal vector of the supporting hyperplane. Such a vector $\mathbf{g}$ is referred to as a *subgradient* at $\mathbf{x}$ and (14.2.1) is known as the *subgradient inequality* and is the generalization of (9.4.4) to non-smooth functions. The set of all subgradients at $\mathbf{x}$ is known as the *subdifferential* at $\mathbf{x}$ and is defined by

$$\partial f(\mathbf{x}) \triangleq \{\mathbf{g}\colon f(\mathbf{x}+\boldsymbol{\delta}) \geqslant f(\mathbf{x}) + \boldsymbol{\delta}^{\mathrm{T}}\mathbf{g} \qquad \forall\, \mathbf{x}+\boldsymbol{\delta} \in K\}. \tag{14.2.2}$$

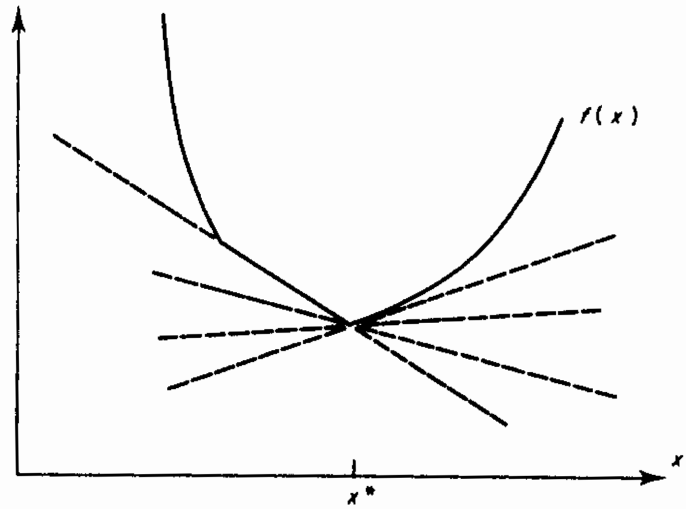

Figure 14.2.1 Supporting hyperplanes to a non-smooth convex function

The notation $\partial f^{(k)} = \partial f(\mathbf{x}^{(k)})$ analogous to $f^{(k)} = f(\mathbf{x}^{(k)})$, etc., is used. It is easy to see that $\partial f(\mathbf{x})$ is a closed convex set, which by Lemma 14.2.1 below is also bounded and therefore compact. In fact a more general result can be established.

Lemma 14.2.1

$\partial f(\mathbf{x})$ is bounded for all $\mathbf{x} \in B \subset$ interior (K) where B is compact.

Proof

If the lemma is false, $\exists$ a sequence $\mathbf{g}^{(k)} \in \partial f(\mathbf{x}^{(k)})$, $\mathbf{x}^{(k)} \in B$, such that $\|\mathbf{g}^{(k)}\|_2 \to \infty$. By compactness $\exists$ a subsequence $\mathbf{x}^{(k)} \to \mathbf{x}'$. Define $\boldsymbol{\delta}^{(k)} = \mathbf{g}^{(k)} / \|\mathbf{g}^{(k)}\|_2^2$. Then $\mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)} \in K$ for k sufficiently large, so by (14.2.1)

$$f(\mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}) \geqslant f^{(k)} + \mathbf{g}^{(k)\mathrm{T}} \boldsymbol{\delta}^{(k)} = f^{(k)} + 1.$$

But in the limit, $f^{(k)} \to f'$, $\boldsymbol{\delta}^{(k)} \to \mathbf{0}$, and so $f(\mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}) \to f'$, which is a contradiction, so the lemma is true. □

If f is differentiable at $\mathbf{x}$, then

$$f(\mathbf{x} + \boldsymbol{\delta}) = f(\mathbf{x}) + \boldsymbol{\delta}^{\mathrm{T}} \nabla f(\mathbf{x}) + o(\|\boldsymbol{\delta}\|)$$

and subtracting this from (14.2.1) gives

$$\boldsymbol{\delta}^{\mathrm{T}}(\mathbf{g} - \nabla f(\mathbf{x})) \leqslant o(\|\boldsymbol{\delta}\|).$$

Choosing $\boldsymbol{\delta} = \theta(\mathbf{g} - \nabla f(\mathbf{x}))$, $\theta \downarrow 0$, shows that $\mathbf{g} = \nabla f$. Hence in this case $\partial f(\mathbf{x})$ is the single vector $\nabla f(\mathbf{x})$.

In Section 14.1 the particular class of polyhedral convex functions $h(\mathbf{c})$, $\mathbb{R}^m \to \mathbb{R}^1$, is defined by

$$h(\mathbf{c}) = \max_i \mathbf{c}^{\mathrm{T}} \mathbf{h}_i + b_i \tag{14.2.3}$$

where $\mathbf{h}_i$ are the columns of a given finite matrix $\mathbf{H}$. Defining

$$\mathscr{A} = \mathscr{A}(\mathbf{c}) \triangleq \{i\colon \mathbf{c}^{\mathrm{T}}\mathbf{h}_i + b_i = h(\mathbf{c})\} \tag{14.2.4}$$

as the set of supporting planes which are active at $\mathbf{c}$ and so attain the maximum, then it is clear that these planes determine the subdifferential $\partial h(\mathbf{c})$. This is proved as follows.

Lemma 14.2.2

$$\partial h(\mathbf{c}) = \operatorname*{conv}_{i\in\mathscr{A}(\mathbf{c})} \mathbf{h}_i \tag{14.2.5}$$

Proof

$\partial h(\mathbf{c})$ is defined by

$$\partial h(\mathbf{c}) = \{\boldsymbol{\lambda}\colon h(\mathbf{c}+\boldsymbol{\delta}) \geqslant h(\mathbf{c}) + \boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{\lambda} \qquad \forall \boldsymbol{\delta}\}. \tag{14.2.6}$$

Let $\boldsymbol{\lambda} = \mathbf{H}\boldsymbol{\mu} \in (14.2.5)$ where $\mu_i \geqslant 0$, $\sum \mu_i = 1$. Then for all $\boldsymbol{\delta}$,

$$\begin{aligned} h(\mathbf{c}) + \boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{\lambda} &= \max_i (\mathbf{c}^{\mathrm{T}}\mathbf{h}_i + b_i) + \sum_{i\in\mathscr{A}} \boldsymbol{\delta}^{\mathrm{T}}\mathbf{h}_i\mu_i \\ &\leqslant \max_{i\in\mathscr{A}} (\mathbf{c}^{\mathrm{T}}\mathbf{h}_i + b_i) + \max_{i\in\mathscr{A}} \boldsymbol{\delta}^{\mathrm{T}}\mathbf{h}_i \\ &= \max_{i\in\mathscr{A}} ((\mathbf{c}+\boldsymbol{\delta})^{\mathrm{T}}\mathbf{h}_i + b_i) \leqslant h(\mathbf{c}+\boldsymbol{\delta}). \end{aligned}$$

Thus $\boldsymbol{\lambda}\in(14.2.6)$. Now let $\boldsymbol{\lambda}\in(14.2.6)$ and assume $\boldsymbol{\lambda}\notin(14.2.5)$. Then by Lemma 14.2.3 below, $\exists \mathbf{s} \neq \mathbf{0}$ such that $\mathbf{s}^{\mathrm{T}}\boldsymbol{\lambda} > \mathbf{s}^{\mathrm{T}}\boldsymbol{\mu}$ $\forall \boldsymbol{\mu}\in(14.2.5)$. Taking $\boldsymbol{\delta} = \alpha\mathbf{s}$, and since $\mathbf{h}_i \in (14.2.5)$ $\forall i \in \mathscr{A}$,

$$\begin{aligned} h(\mathbf{c}) + \boldsymbol{\delta}^{\mathrm{T}}\boldsymbol{\lambda} &= \max_i (\mathbf{c}^{\mathrm{T}}\mathbf{h}_i + b_i) + \alpha\mathbf{s}^{\mathrm{T}}\boldsymbol{\lambda} \\ &> \mathbf{c}^{\mathrm{T}}\mathbf{h}_i + b_i + \alpha\mathbf{s}^{\mathrm{T}}\mathbf{h}_i \qquad \forall i \in \mathscr{A} \\ &= \max_{i\in\mathscr{A}} ((\mathbf{c}+\alpha\mathbf{s})^{\mathrm{T}}\mathbf{h}_i + b_i) \\ &\geqslant \max_i ((\mathbf{c}+\alpha\mathbf{s})^{\mathrm{T}}\mathbf{h}_i + b_i) = h(\mathbf{c}+\boldsymbol{\delta}) \end{aligned}$$

for α sufficiently small, since the max is then achieved on a subset of $\mathscr{A}$. Thus (14.2.6) is contradicted, proving $\boldsymbol{\lambda}\in(14.2.5)$. Hence the definitions of $\partial h(\mathbf{c})$ in (14.2.5) and (14.2.6) are equivalent. $\square$

Examples of this result for a number of particular cases of $h(\mathbf{c})$ and $\mathbf{H}$ are given in more detail in Section 14.1. For convex functions involving a norm, an alternative description of $\partial h(\mathbf{c})$ is provided by (14.3.7) and (14.3.8). The above lemma makes use of the following important result.

Lemma 14.2.3 (Separating hyperplane lemma for convex sets)

If K is a closed convex set and $\boldsymbol{\lambda} \notin K$ then there exists a hyperplane which separates $\boldsymbol{\lambda}$ and K (see Figure 14.2.2).

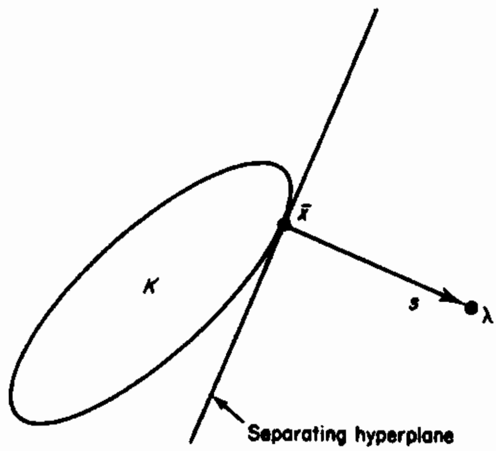

Figure 14.2.2 Existence of a separating hyperplane

Proof

Let $\mathbf{x}_0 \in K$. Then the set $\{\mathbf{x}: \|\mathbf{x}-\boldsymbol{\lambda}\|_2 \leqslant \|\mathbf{x}_0-\boldsymbol{\lambda}\|_2\}$ is bounded and so there exists a minimizer, $\bar{\mathbf{x}}$ say, to the problem: $\min \|\mathbf{x}-\boldsymbol{\lambda}\|_2$, $\mathbf{x}\in K$. Then for any $\mathbf{x}\in K$, $\theta\in[0,1]$,

$$\|(1-\theta)\bar{\mathbf{x}}+\theta\mathbf{x}-\boldsymbol{\lambda}\|_2^2 \geqslant \|\bar{\mathbf{x}}-\boldsymbol{\lambda}\|_2^2$$

and in the limit $\theta\downarrow 0$ it follows that

$$(\mathbf{x}-\bar{\mathbf{x}})^{\mathrm{T}}(\boldsymbol{\lambda}-\bar{\mathbf{x}}) \leqslant 0 \qquad \forall \mathbf{x}\in K.$$

Thus the vector $\mathbf{s}=\boldsymbol{\lambda}-\bar{\mathbf{x}}\neq\mathbf{0}$ satisfies both $\mathbf{s}^{\mathrm{T}}(\boldsymbol{\lambda}-\bar{\mathbf{x}})>0$ and $\mathbf{s}^{\mathrm{T}}(\mathbf{x}-\bar{\mathbf{x}})\leqslant 0$ $\forall \mathbf{x}\in K$ and hence

$$\mathbf{s}^{\mathrm{T}}\boldsymbol{\lambda} > \mathbf{s}^{\mathrm{T}}\mathbf{x} \qquad \forall \mathbf{x}\in K. \tag{14.2.7}$$

The hyperplane $\mathbf{s}^{\mathrm{T}}(\mathbf{x}-\bar{\mathbf{x}})=0$ thus separates K and $\boldsymbol{\lambda}$ as illustrated in Figure 14.2.2 □

At any point $\mathbf{x}'$ at which ∇f does not exist, it nonetheless happens that the directional derivative of the convex function in any direction is well defined. It is again assumed that $f(\mathbf{x})$ is defined on a convex set $K\subset\mathbb{R}^n$ and $\mathbf{x}'\in$ interior (K). The following preliminary result is required.

Lemma 14.2.4

Let $\mathbf{x}^{(k)}\to\mathbf{x}'$ and $\mathbf{g}^{(k)}\in\partial f^{(k)}$. Then any accumulation point of $\{\mathbf{g}^{(k)}\}$ is in $\partial f'$.

Proof

For any $\mathbf{y}\in K$, (14.2.1) can be written as

$$f(\mathbf{y})\geqslant f^{(k)}+(\mathbf{y}-\mathbf{x}^{(k)})^{\mathrm{T}}\mathbf{g}^{(k)}.$$

Taking any subsequence for which $\mathbf{g}^{(k)} \to \mathbf{g}'$ it follows that

$$f(\mathbf{y}) \geqslant f' + (\mathbf{y} - \mathbf{x}')^{\mathrm{T}}\mathbf{g}' \qquad \forall \mathbf{y} \in K,$$

that is $\mathbf{g}' \in \partial f'$. □

A result for the directional derivative at $\mathbf{x}'$ in a direction $\mathbf{s}$ can now be given in a quite general way for any directional sequence (see Section 9.2).

Lemma 14.2.5

If $\mathbf{x}^{(k)} \to \mathbf{x}'$ *is any directional sequence with* $\delta^{(k)} > 0$ *such that* $\delta^{(k)} \to 0$ *and* $\mathbf{s}^{(k)} \to \mathbf{s}$ *in* (9.2.1) (*and allowing* $\mathbf{s} = \mathbf{0}$), *then*

$$\lim_{k \to \infty} \frac{f^{(k)} - f'}{\delta^{(k)}} = \max_{\mathbf{g} \in \partial f'} \mathbf{s}^{\mathrm{T}}\mathbf{g}. \tag{14.2.8}$$

Proof

If $\mathbf{g}^{(k)} \in \partial f^{(k)}$ then it follows from (14.2.1) that for k sufficiently large

$$f' \geqslant f^{(k)} - \delta^{(k)}\mathbf{s}^{(k)\mathrm{T}}\mathbf{g}^{(k)}$$

and

$$f^{(k)} \geqslant f' + \delta^{(k)}\mathbf{s}^{(k)\mathrm{T}}\mathbf{g} \qquad \forall \mathbf{g} \in \partial f'$$

both hold, and hence

$$\mathbf{s}^{(k)\mathrm{T}}\mathbf{g}^{(k)} \geqslant \frac{f^{(k)} - f'}{\delta^{(k)}} \geqslant \max_{\mathbf{g} \in \partial f'} \mathbf{s}^{\mathrm{T}}\mathbf{g}. \tag{14.2.9}$$

Since $\partial f^{(k)}$ is bounded in a neighbourhood of $\mathbf{x}'$ (Lemma 14.2.1), there exists a subsequence for which $\mathbf{g}^{(k)} \to \mathbf{g}'$, and $\mathbf{g}' \in \partial f'$ by Lemma 14.2.4. If (14.2.8) is not true then (14.2.9) gives a contradiction in the limit of such a subsequence. □

Thus (14.2.8) shows that the directional derivative is determined by an extreme supporting hyperplane whose subgradient gives the greatest slope: see Figure 14.2.1 for example.

It is possible to deduce from this result that if $\mathbf{x}^*$ is a local minimizer of $f(\mathbf{x})$, then $f^{(k)} \geqslant f^*$ for all k sufficiently large, and hence from (14.2.8) that

$$\max_{\mathbf{g} \in \partial f^*} \mathbf{s}^{\mathrm{T}}\mathbf{g} \geqslant 0 \qquad \forall \mathbf{s} : \|\mathbf{s}\| = 1. \tag{14.2.10}$$

Thus a *first order necessary condition* for a local minimum is that the directional derivative is non-negative in all directions. This can be stated alternatively as

$$\mathbf{0} \in \partial f^* \tag{14.2.11}$$

which generalizes the condition $\mathbf{g}^* = \mathbf{0}$ for smooth functions. Clearly (14.2.11) implies (14.2.10). If $\mathbf{0} \notin \partial f'$ then by Lemma 14.2.3 (with $\boldsymbol{\lambda} = \mathbf{0}$, $K = \partial f'$)

there exists a vector $\mathbf{s} = -\bar{\mathbf{g}}/\|\bar{\mathbf{g}}\|_2$ for which $\mathbf{s}^T\mathbf{g} < 0 \ \forall \ \mathbf{g} \in \partial f'$, where $\bar{\mathbf{g}}$ is the vector which minimizes $\|\mathbf{g}\|_2 \ \forall \ \mathbf{g} \in \partial f'$. Applying this result at $\mathbf{x}^*$ shows that (14.2.10) and (14.2.11) are equivalent. It is now immediate from (14.2.1) that either (14.2.10) or (14.2.11) is also a sufficient condition for a global minimizer at $\mathbf{x}^*$. In fact the vector $\mathbf{s} = -\bar{\mathbf{g}}/\|\bar{\mathbf{g}}\|_2$ defined above is the *steepest descent vector* at $\mathbf{x}'$. Assuming $\mathbf{0} \notin \partial f'$ then from (14.2.8) the direction of least slope is defined by

$$\min_{\|\mathbf{s}\|_2=1} \max_{\mathbf{g}\in\partial f'} \mathbf{s}^T\mathbf{g} = \max_{\mathbf{g}\in\partial f'} \min_{\|\mathbf{s}\|_2=1} \mathbf{s}^T\mathbf{g}$$
$$= \max_{\mathbf{g}\in\partial f'} -\|\mathbf{g}\|_2 = -\|\bar{\mathbf{g}}\|_2 \tag{14.2.12}$$

and hence the least slope is attained when $\mathbf{s} = -\bar{\mathbf{g}}/\|\bar{\mathbf{g}}\|_2$. The justification of interchanging the min and max operations is explored in Question 14.7.

The main aim in introducing the above development is to apply it to composite functions of the form

$$\phi(\mathbf{x}) = f(\mathbf{x}) + h(\mathbf{c}(\mathbf{x})) \tag{14.2.13}$$

where $f(\mathbf{x})$ ($\mathbb{R}^n \to \mathbb{R}^1$) and $\mathbf{c}(\mathbf{x})$ ($\mathbb{R}^n \to \mathbb{R}^m$) are smooth ($\mathbb{C}^1$) functions and $h(\mathbf{c})$ ($\mathbb{R}^m \to \mathbb{R}^1$) is convex but non-smooth ($\mathbb{C}^0$). Let $\mathbf{x}^{(k)} \to \mathbf{x}'$ be a directional sequence with $\delta^{(k)} \downarrow 0$ and $\mathbf{s}^{(k)} \to \mathbf{s}$ in (9.2.1). By Taylor series

$$f^{(k)} = f' + \delta^{(k)}\mathbf{g}'^T\mathbf{s}^{(k)} + o(\delta^{(k)})$$

so $(f^{(k)} - f')/\delta^{(k)} \to \mathbf{g}'^T\mathbf{s}$. Likewise

$$\mathbf{c}^{(k)} = \mathbf{c}' + \delta^{(k)}\mathbf{A}'^T\mathbf{s}^{(k)} + o(\delta^{(k)})$$

so $\mathbf{c}^{(k)} \to \mathbf{c}'$ is a directional sequence in $\mathbb{R}^m$ with $(\mathbf{c}^{(k)} - \mathbf{c}')/\delta^{(k)} \to \mathbf{A}'^T\mathbf{s}$. Thus by applying Lemma 14.2.5 to $h(\mathbf{c})$, it follows that

$$\lim_{k\to\infty} \frac{\phi^{(k)} - \phi'}{\delta^{(k)}} = \max_{\boldsymbol{\lambda}\in\partial h'} \mathbf{s}^T(\mathbf{g}' + \mathbf{A}'\boldsymbol{\lambda}) \tag{14.2.14}$$

and this gives the directional derivative at $\mathbf{x}'$ in the direction $\mathbf{s}$ for the function $\phi(\mathbf{x})$ in (14.2.13). Another more general result about directional derivatives is proved in Lemma 14.5.1 and its corollary.

It can be deduced from (14.2.14) that if $\mathbf{x}^*$ is a local minimizer of $\phi(\mathbf{x})$, then $\phi^{(k)} \geqslant \phi^*$ for all k sufficiently large, and hence

$$\max_{\boldsymbol{\lambda}\in\partial h^*} \mathbf{s}^T(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) \geqslant 0 \qquad \forall \mathbf{s}: \|\mathbf{s}\| = 1. \tag{14.2.15}$$

This is a first order necessary condition for a local minimizer which like (14.2.10) can be interpreted as a non-negative directional derivative in all directions. Again the result can be stated alternatively as

$$\mathbf{0} \in \partial\phi(\mathbf{x}^*) \triangleq \{\boldsymbol{\gamma}: \boldsymbol{\gamma} = \mathbf{g} + \mathbf{A}\boldsymbol{\lambda} \quad \forall \boldsymbol{\lambda} \in \partial h\}_{\mathbf{x}=\mathbf{x}^*}. \tag{14.2.16}$$

The set $\partial\phi^*$ thus defined, although convex and compact, is not the subdifferential because ϕ may not be a convex function, but it is convenient to use the same notation. (It is in fact the *generalized gradient* of Clarke (1975).) Its definition in (14.2.16) is in the nature of a *generalized chain rule*, by analogy with the expression $\nabla\phi = \mathbf{g} + \mathbf{A}\nabla h$ for smooth functions. The equivalence of (14.2.15) and (14.2.16) is again a consequence of the existence of a separating hyperplane (Lemma 14.2.3). Yet another way to state this condition is to introduce the Lagrangian function

$$\mathscr{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{c}(\mathbf{x}) \tag{14.2.17}$$

Then an equivalent statement of (14.2.16) is as follows.

Theorem 14.2.1 (First order necessary conditions)

If $\mathbf{x}^*$ *minimizes* $\phi(\mathbf{x})$ *in* (14.2.3) *then there exists a vector* $\boldsymbol{\lambda}^* \in \partial h^*$ *such that*

$$\nabla\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}. \tag{14.2.18}$$

Proof

Immediate since $\partial\phi^*$ is the set of vectors $\nabla\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda})$ for all $\boldsymbol{\lambda} \in \partial h^*$. □

This form illustrates the close relationship between the vector $\boldsymbol{\lambda}^* \in \partial h^*$ and the Lagrange multipliers in Theorem 9.1.1. These conditions are illustrated by Questions 14.8, 14.9, 14.10, and 14.12. In general, since $\phi(\mathbf{x})$ may not be convex, the conditions of Theorem 14.2.1 are not sufficient.

In view of this last observation it is important to consider second order conditions for $\mathbf{x}^*$ to be a minimizer of the composite function (14.2.13). These conditions again exhibit a close relationship with the results in Section 9.3. The approach is based on that taken by Fletcher and Watson (1980). In considering second order conditions it is necessary to restrict the possible directions to those having zero directional derivative so that second order effects become important. This is illustrated by the contours in Figure 14.2.3. $\mathbf{x}^*$ satisfies first order conditions but there are directions of zero slope along the derivative discontinuity $c_1(\mathbf{x}) = 0$ (broken line), so that first order conditions are not sufficient. However sufficient second order conditions can be derived which imply that $\mathbf{x}^*$ is a local minimizer—see Question 14.9. In general let $\boldsymbol{\lambda}^*$ be any vector which exists in Theorem 14.2.1 and consider the set

$$X = \{\mathbf{x}\colon h(\mathbf{c}(\mathbf{x})) = h(\mathbf{c}(\mathbf{x}^*)) + (\mathbf{c}(\mathbf{x}) - \mathbf{c}(\mathbf{x}^*))^{\mathrm{T}}\boldsymbol{\lambda}^*\}. \tag{14.2.19}$$

Define $\mathscr{G}^*$ as the set of normalized feasible directions with respect to X at $\mathbf{x}^*$. (That is $\mathbf{s} \in \mathscr{G}^*$ implies that there exists a directional sequence $\mathbf{x}^{(k)} \to \mathbf{x}^*$, feasible in (14.2.19), such that $\mathbf{s}^{(k)} \to \mathbf{s}$ in (9.2.1) with $\sigma = 1$.) $\mathscr{G}^*$ can be regarded as the set of stationary directions for $h(\mathbf{c}(\mathbf{x}))$, having linearized $h(\mathbf{c})$ but not $\mathbf{c}(\mathbf{x})$. It is possible to show that these directions are closely related to the set G^* of normalized directions of zero slope, that is

$$G^* \triangleq \left\{\mathbf{s}\colon \max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) = 0,\ \ \|\mathbf{s}\|_2 = 1\right\}. \tag{14.2.20}$$

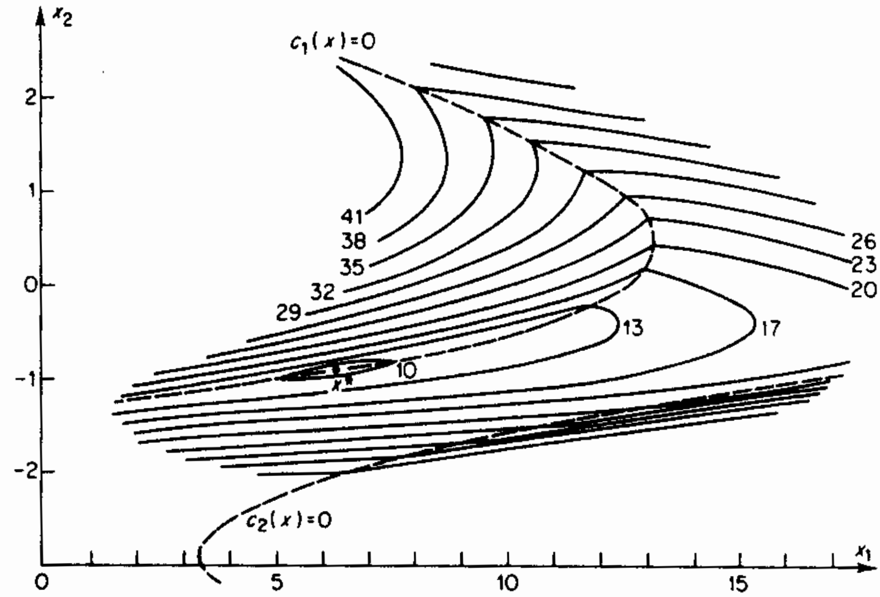

Figure 14.2.3 Contours for the scaled L_1 Freudenstein and Roth problem, Question 14.9

G^* can be regarded as the set of stationary directions for $h(\mathbf{c}(\mathbf{x}))$ having linearized both $h(\mathbf{c})$ and $\mathbf{c}(\mathbf{x})$. The extent to which $\mathscr{G}^*$ and G^* correspond is important and it is first shown that $\mathscr{G}^*$ is a subset of G^*.

Lemma 14.2.6

$\mathscr{G}^* \subset G^*$.

Proof

Let $\mathbf{s} \in \mathscr{G}^*$ so a directional sequence in X exists with $\mathbf{s}^{(k)} \to \mathbf{s}$, $\|\mathbf{s}\|_2 = 1$. Using (14.2.14), (14.2.13), (14.2.19), and a Taylor series it follows that

$$\max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) = \lim_{k \to \infty} \frac{\phi^{(k)} - \phi^*}{\delta^{(k)}}$$

$$= \lim_{k \to \infty} \frac{f^{(k)} - f^* + h(\mathbf{c}^{(k)}) - h(\mathbf{c}^*)}{\delta^{(k)}}$$

$$= \lim_{k \to \infty} \frac{f^{(k)} - f^* + (\mathbf{c}^{(k)} - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda}^*}{\delta^{(k)}}$$

$$= \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^*).$$

It then follows from (14.2.18) that $\mathbf{s} \in G^*$. $\square$

To give a general result going the other way is not always possible although it can be done in important special cases. Further discussion of this is given later in this section: at present the regularity assumption

$$\mathscr{G}^* = G^* \tag{14.2.21}$$

is made (which depends on $\boldsymbol{\lambda}^*$ if more than one such vector exists). It is now possible to state the second order conditions. In doing this it is assumed that f and $\mathbf{c}$ are $\mathbb{C}^2$ functions. Note that as usual the regularity assumption is needed only in the necessary conditions.

Theorem 14.2.2 (Second order necessary conditions)

If $\mathbf{x}^*$ *minimizes* $\phi(\mathbf{x})$ *then Theorem* 14.2.1 *holds; for each vector* $\boldsymbol{\lambda}^*$ *which thus exists, if* (14.2.21) *holds, then*

$$\mathbf{s}^{\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{s} \geqslant 0 \qquad \forall \mathbf{s} \in G^*. \tag{14.2.22}$$

Proof

For any $\mathbf{s} \in G^*$, $\mathbf{s} \in \mathscr{G}^*$ by (14.2.21) and hence $\exists$ a directional sequence $\mathbf{x}^{(k)} \to \mathbf{x}^*$, feasible in (14.2.19). A Taylor expansion of $\mathscr{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ about $\mathbf{x}^*$ yields

$$\begin{aligned}\mathscr{L}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^*) &= f^* + \mathbf{c}^{*\mathrm{T}}\boldsymbol{\lambda}^* + \mathbf{e}^{(k)\mathrm{T}}\nabla\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) \\ &\quad + \tfrac{1}{2}\mathbf{e}^{(k)\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{e}^{(k)} + o(\|\mathbf{e}^{(k)}\|^2)\end{aligned}$$

where $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*$, and using (14.2.18) and (9.2.1),

$$\mathscr{L}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^*) = f^* + \mathbf{c}^{*\mathrm{T}}\boldsymbol{\lambda}^* + \tfrac{1}{2}\delta^{(k)^2}\mathbf{s}^{(k)\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{s}^{(k)} + o(\delta^{(k)^2}). \tag{14.2.23}$$

Since $\mathbf{x}^{(k)}$ is feasible in (14.2.19) it follows from (14.2.13) that

$$\phi^{(k)} = \phi^* + \tfrac{1}{2}\delta^{(k)^2}\mathbf{s}^{(k)\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{s}^{(k)} + o(\delta^{(k)^2}).$$

Since $\mathbf{x}^*$ is a local minimizer, $\phi^{(k)} \geqslant \phi^*$ for all k sufficiently large, so dividing by $\frac{1}{2}\delta^{(k)^2}$ and taking the limit yields (14.2.22). □

For sufficient conditions, it is firstly observed that if the directional derivative is positive in all directions, that is

$$\max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) > 0 \qquad \forall \mathbf{s}: \|\mathbf{s}\| = 1,$$

or equivalently G^* is empty, then the conditions of Theorem 14.2.1 imply that $\mathbf{x}^*$ is a strict and isolated local minimizer (an example is given in Question 14.12). This result is in fact a special case of Theorem 14.2.3 below. When G^* is non-empty, second order effects come into play and this is expressed in the following.

Theorem 14.2.3 (Second order sufficient conditions)

If there exists $\boldsymbol{\lambda}^* \in \partial h^*$ *such that* (14.2.18) *holds, and if*

$$\mathbf{s}^{\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{s} > 0 \qquad \forall \mathbf{s} \in G^* \tag{14.2.24}$$

then $\mathbf{x}^*$ *is a strict and isolated local minimizer of* $\phi(\mathbf{x})$.

Proof

Assume that $\mathbf{x}^*$ is not strict, so that $\exists$ a sequence and hence a directional sequence $\mathbf{x}^{(k)} \to \mathbf{x}^*$ such that $\phi^{(k)} \leqslant \phi^*$. By (14.2.18),

$$0 \leqslant \max_{\lambda \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) = \mu$$

say. If $\mu > 0$ then $\lim (\phi^{(k)} - \phi^*)/\delta^{(k)} = \mu$ (see (14.2.14)) which contradicts $\phi^{(k)} \leqslant \phi^*$. Thus $\mu = 0$ and hence $\mathbf{s} \in G^*$. Now from (14.2.17) and (14.2.13)

$$\begin{aligned} \mathscr{L}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^*) - \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) &= f^{(k)} + \mathbf{c}^{(k)\mathrm{T}}\boldsymbol{\lambda}^* - f^* - \mathbf{c}^{*\mathrm{T}}\boldsymbol{\lambda}^* \\ &= \phi^{(k)} - \phi^* - (h(\mathbf{c}^{(k)}) - h(\mathbf{c}^*) - (\mathbf{c}^{(k)} - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda}^*) \\ &\leqslant \phi^{(k)} - \phi^* \end{aligned}$$

using the subgradient inequality. Hence from (14.2.23)

$$0 \geqslant \phi^{(k)} - \phi^* \geqslant \tfrac{1}{2}\delta^{(k)^2}\mathbf{s}^{(k)\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{s}^{(k)} + o(\delta^{(k)^2}).$$

Dividing by $\frac{1}{2}\delta^{(k)^2}$ and taking the limit contradicts (14.2.24) and establishes that $\mathbf{x}^*$ is a strict minimizer.

Now assume that $\mathbf{x}^*$ is not isolated so that $\exists$ a sequence of local solutions $\mathbf{x}^{(k)} \to \mathbf{x}^*$. A directional subsequence can therefore be extracted for which $\mathbf{s}^{(k)} \to \mathbf{s}$ in (9.2.1) with $\|\mathbf{s}\|_2 = 1$, and for which

$$\exists\, \boldsymbol{\lambda}^{(k)} \in \partial h^{(k)}, \qquad \mathbf{g}^{(k)} + \mathbf{A}^{(k)}\boldsymbol{\lambda}^{(k)} = \mathbf{0}, \qquad \boldsymbol{\lambda}^{(k)} \to \boldsymbol{\lambda}^*$$

by virtue of Theorem 14.2.1 at $\mathbf{x}^{(k)}$, and Lemma 14.2.4. By taking Taylor expansions about $\mathbf{x}^{(k)}$ and using the subgradient inequality

$$\begin{aligned} \phi^* &= \phi(\mathbf{x}^{(k)} - \boldsymbol{\delta}^{(k)}) + h(\mathbf{c}(\mathbf{x}^{(k)} - \boldsymbol{\delta}^{(k)})) \\ &\geqslant \phi^{(k)} - \boldsymbol{\delta}^{(k)\mathrm{T}}(\mathbf{g}^{(k)} + \mathbf{A}^{(k)}\boldsymbol{\lambda}^{(k)}) + o(\|\boldsymbol{\delta}^{(k)}\|) \\ &= \phi^{(k)} + o(\|\boldsymbol{\delta}^{(k)}\|) \end{aligned}$$

by $\mathbf{g}^{(k)} + \mathbf{A}^{(k)}\boldsymbol{\lambda}^{(k)} = \mathbf{0}$. Define μ as in the first part of the proof: this result provides a contradiction if $\mu > 0$. Hence $\mu = 0$ and $\mathbf{s} \in G^*$. Taylor expansions for $\mathbf{g}^{(k)}$ and $\mathbf{A}^{(k)}$ and $\boldsymbol{\lambda}^{(k)} \to \boldsymbol{\lambda}^*$ give

$$\mathbf{g}^{(k)} + \mathbf{A}^{(k)}\boldsymbol{\lambda}^{(k)} = \mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* + \nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\boldsymbol{\delta}^{(k)} + o(\|\boldsymbol{\delta}^{(k)}\|).$$

It follows that $\boldsymbol{\delta}^{(k)\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\boldsymbol{\delta}^{(k)} + o(\|\boldsymbol{\delta}^{(k)}\|^2) = 0$ and hence, in the limit, $\mathbf{s}^{\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)\mathbf{s} = 0$ which contradicts (14.2.24). Thus $\mathbf{x}^*$ is also an isolated minimizer. $\square$

These conditions (with G^* non-empty) are illustrated by Questions 14.9 and 14.10.

The second order conditions for nonlinear programming given in Section 9.3 are almost necessary and sufficient because the regularity assumption is mild and there is a gap only in the case of zero curvature. The same is *not* true here and it is important to realize that there are well-behaved cases in which $\mathscr{G}^* \neq G^*$. These cases arise when the linearization of $h(\mathbf{c})$ in (14.2.19) is not valid; for example when $h(\mathbf{c})$ is smooth but nonlinear. A property which enables further progress to be made is the following. If there exists an open neighbourhood Ω of $\mathbf{c}^*$ such that

$$h(\mathbf{c}) = h(\mathbf{c}^*) + \max_{\boldsymbol{\lambda}\in\partial h^*} (\mathbf{c} - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda} \tag{14.2.29}$$

for all $\mathbf{c}\in\Omega$, then $h(\mathbf{c})$ is referred to as being *locally linear* at $\mathbf{c}^*$. This property is clearly true for any polyhedral convex function but also enables certain exact penalty function problems with smooth norms to be handled, as described in Lemma 14.3.1. Another construction is also required. Denote the dimension of $\partial h(\mathbf{c}^*)$ by l^* ($l^* \leqslant m$) and let the matrix $\mathbf{D}^*\in\mathbb{R}^{m\times l^*}$ have columns $\mathbf{d}_i^*$ $i = 1, 2, \ldots, l^*$ which provide a basis for $\partial h(\mathbf{c}^*) - \boldsymbol{\lambda}^*$. That is to say, $\partial h(\mathbf{c}^*)$ can be expressed as

$$\partial h(\mathbf{c}^*) = \{\boldsymbol{\lambda}\colon \boldsymbol{\lambda} = \boldsymbol{\lambda}^* + \mathbf{D}^*\mathbf{u},\ \mathbf{u}\in U^* \subset \mathbb{R}^{l^*}\} \tag{14.2.30}$$

where U^*, which is the image of $\partial h(\mathbf{c}^*) - \boldsymbol{\lambda}^*$ is also convex and compact. The idea is essentially that contained in Osborne's (1985) introduction of *structure functionals* for polyhedral convex functions, taken up more generally by Womersley (1984b). Note that $\boldsymbol{\lambda}^*$ is essentially an arbitrary vector in $\partial h(\mathbf{c}^*)$: the choice of the vector $\boldsymbol{\lambda}^*$ as defined in (14.2.18) is made for reasons of convenience. A mild independence assumption now enables (14.2.21) to be established when $h(\mathbf{c})$ is locally linear at $\mathbf{c}^*$.

Lemma 14.2.7 (Sufficient conditions for regularity)

If $\mathbf{x}^*$ *satisfies first order conditions* (14.2.18), *if* $h(\mathbf{c})$ *is locally linear at* $\mathbf{c}^*$, *and if*

$$\operatorname{rank}(\mathbf{A}^*\mathbf{D}^*) = l^* \tag{14.2.31}$$

then $\mathscr{G}^* = G^*$.

Proof

Because $G^* \supseteq \mathscr{G}^*$ it is sufficient to prove the reverse inclusion. Let $\mathbf{s}\in G^*$ and define

$$\partial h_s^* = \{\boldsymbol{\lambda}\in\partial h^*\colon \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) = 0\}$$

observing that ∂h_s^* depends on $\mathbf{s}$. It follows from (14.2.18) that

$$\mathbf{s}^{\mathrm{T}}\mathbf{A}^*(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) = 0 \qquad \forall\, \boldsymbol{\lambda}\in\partial h_s^*$$

and from (14.2.20) that

$$G^* = \left\{\mathbf{s}\colon\ \|\mathbf{s}\|_2 = 1,\quad \max_{\boldsymbol{\lambda}\in\partial h^*} \mathbf{s}^{\mathrm{T}}\mathbf{A}^*(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) = 0\right\} \tag{14.2.32}$$

and hence that

$$\mathbf{s}^{\mathrm{T}}\mathbf{A}^*(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) < 0 \qquad \forall\, \boldsymbol{\lambda}\in\partial h^*\backslash\partial h_s^*.$$

Let the dimension of ∂h_s^* be l_s^* ($l_s^* \leqslant l^*$) and assume without loss of generality that the vectors $\mathbf{d}_i^*$ $i = 1, 2, \ldots, l_s^*$ form a basis for $\partial h_s^* - \boldsymbol{\lambda}^*$. Hence

$$\mathbf{s}^{\mathrm{T}}\mathbf{A}^*\mathbf{d}_i^* \begin{cases} = 0 & i = 1, 2, \ldots, l_s^* \\ < 0 & i = l_s^* + 1, \ldots, l^*. \end{cases}$$

If $l_s^* = n$ then it follows that $\mathbf{s}^{\mathrm{T}}\mathbf{A}^*\mathbf{D}^* = \mathbf{0}^{\mathrm{T}}$ and hence from the rank assumption it follows that $\mathbf{s} = \mathbf{0}$, which contradicts $\mathbf{s} \in G^*$. For $l_s^* < n$ it is possible to construct a smooth arc $\mathbf{x}(\theta)$ $\theta \in [0, \bar{\theta})$ for which $\mathbf{x}(0) = \mathbf{x}^*$ and $\dot{\mathbf{x}}(0) = \mathbf{s}$ in the manner of Lemma 9.2.2. (The appropriate nonlinear equation in (9.2.3) is $r_i(\mathbf{x}, \theta) = (\mathbf{c}(\mathbf{x}) - \mathbf{c}^*)^{\mathrm{T}}\mathbf{d}_i^* - \theta\mathbf{s}^{\mathrm{T}}\mathbf{A}^*\mathbf{d}_i^* = 0$ and the Jacobian matrix is $[\mathbf{A}^*\mathbf{D}^* \vdots \mathbf{B}]$ where $\mathbf{B}$ is $nx(n - l^*)$.) It therefore follows that

$$(\mathbf{c}(\mathbf{x}(\theta)) - \mathbf{c}^*)^{\mathrm{T}}\mathbf{d}_i^* = \theta\mathbf{s}^{\mathrm{T}}\mathbf{A}^*\mathbf{d}_i \begin{cases} = 0 & i = 1, 2, \ldots, l_s^* \\ < 0 & i = l_s^* + 1, \ldots, l^* \end{cases}$$

and hence that

$$(\mathbf{c}(\mathbf{x}(\theta)) - \mathbf{c}^*)^{\mathrm{T}}(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) \begin{cases} = 0 & \boldsymbol{\lambda} \in \partial h_s^* \\ < 0 & \boldsymbol{\lambda} \in \partial h^* \backslash \partial h_s^* \end{cases}$$

or equivalently that

$$\max_{\boldsymbol{\lambda} \in \partial h^*} (\mathbf{c}(\mathbf{x}(\theta)) - \mathbf{c}^*)^{\mathrm{T}}(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) = 0.$$

Finally using (14.2.29) it follows that there exists a neighbourhood of $\mathbf{c}^*$ such that

$$h(\mathbf{c}(\mathbf{x}(\theta))) = h(\mathbf{c}^*) + (\mathbf{c}(\mathbf{x}(\theta)) - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda}^*$$

and taking any sequence $\theta^{(k)} \downarrow 0$ gives a directional sequence that is feasible in (14.2.19), so that $\mathbf{s} \in \mathscr{G}^*$. $\square$

Of course these are not the only sufficient conditions for $\mathscr{G}^* = G^*$, for example the same result holds if $\mathbf{c}(\mathbf{x})$ is affine (in place of (14.2.31)) or if it can be shown that G^* is empty. However, it does have the advantage that the rank condition (14.2.31) is readily checked when $\mathbf{c}(\mathbf{x})$ is nonlinear. Basis vectors for the set ∂h^* are readily determined in all practical cases, including the functions $h(\mathbf{c})$ listed in (14.1.3). For the max function (I), if $\mathscr{A}^* = \{i : c_i^* = h^*\}$ is the index set over which the max is attained, and $q \in \mathscr{A}^*$ is arbitrary, then $\mathbf{e}_j - \mathbf{e}_q$ $j \in \mathscr{A}^* \backslash q$ is a suitable basis set for the vectors $\mathbf{d}_i^*$. For the function $h(\mathbf{c}) = \|\mathbf{c}^+\|_\infty$ (II), if $h^* > 0$ then the same basis set as for (I) is valid. If $h^* = 0$ then $\mathscr{A}^*$ is the same and $\mathbf{e}_j^*$ $j \in \mathscr{A}^*$ is a basis set. For $h(\mathbf{c}) = \|\mathbf{c}\|_\infty$ (III) define $\mathscr{A}^* = \{j : |c_j^*| = h^*\}$. If $h^* > 0$ then a basis set is $\operatorname{sign}(c_j^*)\mathbf{e}_j - \operatorname{sign}(c_q^*)\mathbf{e}_q$ $j \in \mathscr{A}^* \backslash q$ where $q \in \mathscr{A}^*$ is arbitrary. If $h^* = 0$ (which implies $\mathbf{c}^* = \mathbf{0}$) then $\mathbf{d}_j^* = \mathbf{e}_j^*$ $j = 1, 2, \ldots, m$ is suitable.

In the case of functions based on the L_1 norm (IV and V) it is appropriate to define $\mathscr{A}^* = \{i : c_i^* = 0\}$. Then a suitable basis in both cases is the set of vectors $\mathbf{e}_i$ $i \in \mathscr{A}^*$. Thus the rank assumption (14.2.31) is equivalent to the assumption that the vectors $\mathbf{a}_i^*$ $i \in \mathscr{A}^*$ are linearly independent which is the standard assumption for regularity in nonlinear programming (Lemma 9.2.2). For the general polyhedral convex function (14.1.2), if $\mathscr{A}^*$ is defined by (14.1.10), then a basis set is any basis for the space spanned by the vectors $\mathbf{h}_j - \mathbf{h}_q$ for all $j \in \mathscr{A}^* - q$ where $q \in \mathscr{A}^*$ is arbitrary. A different example of some interest is $h(\mathbf{c}) = \|\mathbf{c}\|_2$ for which if $h(\mathbf{c}^*) = 0$ then $\mathbf{d}_j = \mathbf{e}_j$ $j = 1, 2, \ldots, m$ is a suitable basis for $\partial h^* - \boldsymbol{\lambda}^*$. Generalizations of this result are given in Lemma 14.3.1.

Another direction set of some interest is the column null space of the matrix $\mathbf{A}^*\mathbf{D}^*$,

$$\mathscr{N}^* = \{\mathbf{s}\colon \|\mathbf{s}\|_2 = 1, \mathbf{s}^T\mathbf{A}^*\mathbf{D}^* = \mathbf{0}^T\} \tag{14.2.33}$$

It follows easily from (14.2.32) and (14.2.30) that $\mathscr{N}^* \subseteq G^*$. In the case that $\boldsymbol{\lambda}^*$ is in the relative interior of ∂h^* (equivalently $\mathbf{0} \in \operatorname{int} U^*$ in (14.2.30)) then *strict complementarity* is said to hold (cf. Section 9.1) and $G^* = \mathscr{N}^*$ can be proved. Clearly if $\mathbf{s} \in G^*$ then

$$\mathbf{s}^T\mathbf{A}^*\mathbf{D}^*\mathbf{u} \leqslant 0 \qquad \forall\, \mathbf{u} \in U^*$$

and because $\mathbf{0} \in \operatorname{int} U^*$ it follows that $\mathbf{s}^T\mathbf{A}^*\mathbf{D}^* = \mathbf{0}^T$. Thus the reverse inclusion is true which proves the equivalence. The relationship between $\mathscr{N}^*$ and G^* is analogous to that between (9.3.4) and (9.3.11) for nonlinear programming.

A few special cases can be summarized. If $h(\mathbf{c})$ is smooth at $\mathbf{c}^*$, then $l^* = 0$ and $\mathscr{N}^* = G^* = \mathbb{R}^n$ whereas usually $\mathscr{G}^* = \varnothing$ (in absence of any locally linear behaviour). If $(\phi^*, \mathbf{x}^*)$ is a regular vertex of the epigraph of $\phi(\mathbf{x})$, then $l^* = n$. If strict complementarity also holds then $\mathscr{N}^* = G^* = \mathscr{G}^* = \varnothing$, the second order conditions (14.2.24) are vacuous and first order information is sufficient to ensure a strict local minimizer. It can also be the case that $(\phi^*, \mathbf{x}^*)$ is a degenerate vertex of the epigraph and $l^* > n$. However, if rank $\mathbf{A}^*\mathbf{D}^* = n$ and strict complementarity holds for some $\boldsymbol{\lambda}^* \in \partial h^*$ then it follows that $\mathscr{N}^* = G^* = \mathscr{G}^* = \varnothing$ as above and again first order information is sufficient.

These results enable a very illuminating comparison to be drawn between the composite nonsmooth optimization problem

$$\text{minimize } f(\mathbf{x}) + h(\mathbf{c}(\mathbf{x})) \tag{14.2.34}$$

and the nonlinear programming problem

$$\begin{aligned} &\text{minimize} \quad f(\mathbf{x}) + \mathbf{c}(\mathbf{x})^T\boldsymbol{\lambda}^* \\ &\text{subject to} \quad \mathbf{D}^{*T}(\mathbf{c}^* - \mathbf{c}(\mathbf{x})) = \mathbf{0}, \end{aligned} \tag{14.2.35}$$

as expressed by the following theorem. In this theorem, multipliers of the constraints in (14.2.35) are denoted by $\mathbf{u}^* \in \mathbb{R}^{l^*}$.

Theorem 14.2.4

If $h(\mathbf{c})$ is locally linear about $\mathbf{c}^*$ then

(i) if $\mathbf{x}^*$ *is a local minimizer of* (14.2.34) *then* $\mathbf{x}^*$ *is both a local solution and a KT point of* (14.2.35),
(ii) *if* $\mathbf{x}^*$ *satisfies second order sufficient conditions for* (14.2.34) *then* $\mathbf{x}^*$ *satisfies second order sufficient conditions for* (14.2.35).

(iii) *If* $\mathbf{x}^*, \mathbf{u}^*$ ($\mathbf{u}^* = \mathbf{0}$) *satisfies second order sufficient conditions for* (14.2.35) *and* $\boldsymbol{\lambda}^*$ *is in the relative interior of* $\partial h(\mathbf{c}^*)$ *then* $\mathbf{x}^*$ *satisfies second order sufficient conditions for* (14.2.34),

(iv) *the condition rank* $(\mathbf{A}^*\mathbf{D}^*) = l^*$ *is sufficient for regularity* ($\mathscr{G}^* = G^*$) *in both cases.*

Proof

In part (i) it follows from Theorem 12.2.1 that $\exists\ \boldsymbol{\lambda}^* \in \partial h(\mathbf{c}^*)$ such that $\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}$. Then (14.2.30) implies that $\mathbf{x}^*$ is a KT point for (14.2.35) with multiplier vector $\mathbf{u}^* = \mathbf{0}$. Let $\mathbf{x}$ be feasible in (14.2.35) so that $\mathbf{D}^{*\mathrm{T}}(\mathbf{c}^* - \mathbf{c}(\mathbf{x})) = \mathbf{0}$, and close enough to $\mathbf{x}^*$ so that (14.2.29) is valid for $\mathbf{c} = \mathbf{c}(\mathbf{x})$. It follows from these that $h(\mathbf{c}(\mathbf{x})) = h(\mathbf{c}^*) + \boldsymbol{\lambda}^{*\mathrm{T}}(\mathbf{c}(\mathbf{x}) - \mathbf{c}^*)$. Because $\mathbf{x}^*$ is a local minimizer of (14.2.34), it follows that $f(\mathbf{x}) + \boldsymbol{\lambda}^{*\mathrm{T}}\mathbf{c}(\mathbf{x}) \geqslant f^* + \boldsymbol{\lambda}^{*\mathrm{T}}\mathbf{c}^*$ and hence $\mathbf{x}^*$ is a local solution of (14.2.35).

For part (ii), second order sufficient conditions for (14.2.34) in Theorem 14.2.3 involve the matrix $\mathbf{W}^* = \nabla^2 f_i^* + \sum \lambda_i^* \nabla^2 c_i^*$. Those for (14.2.35) involve the matrix $\mathbf{W}^* = \nabla^2 f_i^* + \sum \lambda_i^* \nabla^2 c_i^* + \sum u_j^* \nabla^2 (\mathbf{D}^{*\mathrm{T}}\mathbf{c})_j$ which is the same matrix since $\mathbf{u}^* = \mathbf{0}$. Theorem 14.2.3 involves the direction set G^* in (14.2.20). Theorem 9.3 involves the direction set given in (9.3.11) which turns out to be the set $\mathscr{N}^*$ in (14.2.33). Part (ii) follows because $\mathscr{N}^* \subset G^*$ as pointed out after (14.2.33).

For part (iii) it follows that $\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}$, and by assumption $\boldsymbol{\lambda}^* \in \partial h(\mathbf{c}^*)$, so $\mathbf{x}^*$ satisfies first order conditions for (14.2.34). Moreover $\boldsymbol{\lambda}^* \in \text{rel int } \partial h(\mathbf{c}^*)$ (i.e. $\mathbf{0} \in \text{int } U^*$) implies $\mathscr{N}^* = G^*$. Hence second order sufficient conditions hold for (14.2.34) by following a reverse argument to that in the previous paragraph. For part (iv) the rank condition is that in (14.2.31) which is the same assumption that is needed to prove (9.3.12) (cf. Lemma 9.2.2). $\square$

The implications of this result are that if $\boldsymbol{\lambda}^* \in \partial h(\mathbf{c}^*)$ then KT points for (14.2.34) and (14.2.35) are equivalent. Moreover if $h(\mathbf{c})$ is locally linear about $\mathbf{c}^*$ and the rank assumption (14.2.31) holds then regularity holds and the second order conditions are 'almost' necessary and sufficient. Thus, except in limiting cases, problems (14.2.34) and (14.2.35) have identical solutions. The objective function in (14.2.35) is essentially the "smooth" part of (14.2.34), and the $\mathbf{c}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\lambda}^*$ term shifts the multipliers to $\mathbf{u}^* = \mathbf{0}$. The constraints in (14.2.35) are essentially the *structure functionals* given by Osborne (1985) in the polyhedral case and enable the nonsmooth part of (14.2.34) to be replaced by a system of nonlinear equations. It will be seen in the derivation of algorithms in Section 14.4 that this replacement allows us to analyse very readily the asymptotic behaviour of iteration sequences for the SNQP method. It also reinforces, but in a different way, the observation in Section 14.1 that the composite NSO problem is equivalent to a nonlinear programming problem ((14.1.6) in that case).

14.3 EXACT PENALTY FUNCTIONS

One of the most important applications of NSO is in the area of nonlinear programming through the use of an exact penalty function. For the special case of an L_1 exact penalty function, a simple presentation is given in Section 12.3. In this section a quite general treatment of the exact penalty function is given, involving an arbitrary norm. To simplify the presentation, two basic types of nonlinear programming problem are considered, the equality constraint problem

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \quad f(\mathbf{x}) \\ &\text{subject to} \quad \mathbf{c}(\mathbf{x}) = \mathbf{0} \end{aligned} \tag{14.3.1}$$

for which the corresponding exact penalty function problem is

$$\underset{\mathbf{x}}{\text{minimize}} \ \phi(\mathbf{x}) \triangleq \nu f(\mathbf{x}) + \|\mathbf{c}(\mathbf{x})\|, \tag{14.3.2}$$

and the inequality constraint problem

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} \quad f(\mathbf{x}) \\ &\text{subject to} \quad \mathbf{c}(\mathbf{x}) \leqslant \mathbf{0} \end{aligned} \tag{14.3.3}$$

for which the corresponding exact penalty function problem is

$$\underset{\mathbf{x}}{\text{minimize}} \ \phi(\mathbf{x}) \triangleq \nu f(\mathbf{x}) + \|\mathbf{c}(\mathbf{x})^+\| \tag{14.3.4}$$

where c_i^+ denotes $\max(c_i, 0)$. There is no difficulty in generalizing further to the mixed problem (7.1.1) and this can be done within the framework of (14.3.3) by replacing $c_i = 0$ by $c_i \leqslant 0$ and $-c_i \leqslant 0$. Note that in (14.3.3) the inequality is in the reverse direction from that considered in Chapter 9. This is equivalent to a change of sign in the constraint residuals $\mathbf{c}(\mathbf{x})$ which must be kept in mind. These penalty functions are exact in the sense of Section 12.3, so that usually (14.3.1) and (14.3.2) (or (14.3.3) and (14.3.4)) are equivalent in that for sufficiently small ν a local solution of (14.3.1) is a local minimizer of (14.3.2), and *vice versa.* This result is made precise in Theorems 14.3.1 and 14.3.2 below. Practical considerations in choosing ν are discussed in Section 12.3. The conditions under which equivalence holds are usually satisfied in practice but there are examples where the problems are not equivalent. Examples where $\mathbf{x}^*$ solves (14.3.3) and not (14.3.4) are given at the end of Theorem 14.3.1. However, these examples are either pathological or limiting cases and need not greatly concern the user. The alternatively possibility is that a local minimizer $\mathbf{x}^*$ of (14.3.2) may not be feasible in (14.3.1), even though the latter may have a solution. This is the same situation as that illustrated in Figure 6.2.1 and described in Section 6.2: it is an inevitable consequence of the use of any penalty approach as a means of inducing global convergence. To circumvent the difficulty is a global minimization problem and hence generally impracticable. Corresponding advantages are that *best* solutions can be determined when no feasible point exists in (14.3.1) and that the difficulty of finding an initial feasible point is avoided. In practice the most likely unfavourable situation which arises when applying an exact penalty function is that a sequence $\{\mathbf{x}^{(k)}\}$ is calculated such that $\phi^{(k)} \to -\infty$, which is an indication that the calculation should be repeated with a smaller value of ν. It may be necessary to pre-scale the $c_i(\mathbf{x})$ to be of comparable magnitude so that the use of a single penalty parameter ν is reasonable. In this case it can be better to rewrite (14.3.2) for example as

$$\phi(\mathbf{x}) = f(\mathbf{x}) + \|\mathbf{Sc}(\mathbf{x})\|, \tag{14.3.5}$$

where $\mathbf{S} = \operatorname{diag} \sigma_i$ is a diagonal matrix of weights σ_i on each of the constraint functions. There is no difficulty in reorganizing the theory to account for this change.

The main attraction in using (14.3.2) or (14.3.4) is that if holds out the possibility of avoiding the sequential nature of the penalty functions in Sections 12.1 and 12.2 so that only a single unconstrained minimization calculation is required. Unfortunately (even if $\|\cdot\|_2$ is used) (14.3.2) and (14.3.4) are not smooth ($\mathbb{C}^1$) functions so that the many effective techniques for smooth minimization described in Part 1 cannot adequately be used. The study of algorithms for non-smooth problems is a relatively recent development and is described in some detail in Section 14.4 and some of the algorithms can be used here. Another approach (Han, 1977; Coleman and Conn, 1982a, b; Mayne, 1980) is to use an algorithm for nonlinear programming as a means of generating a direction of search, and to use the exact penalty function as the criterion function to be minimized (approximately) in the line search. This approach often works well in practice but unfortunately can fail (see Fletcher, 1981) and I think it is important to take into account the non-smooth nature of the penalty function when choosing the direction in which to search. The discussion of algorithms in Sections 14.4 and 4.5 leads to a globally convergent algorithm for composite NSO problems (Fletcher, 1982a) which works well when applied to minimize exact penalty functions (Fletcher 1981), and it is this type of algorithm which I currently favour. Use of the $\|\cdot\|_1$ is the most convenient for the reasons given in Sections 12.3 and 12.4 where more details are given.

The theory for exact penalty functions can be set out in a very concise and general way. The definition in (14.3.2) allows the use of any norm and in (14.3.4) any *monotonic* norm ($|\mathbf{x}| \leqslant |\mathbf{y}| \Rightarrow \|\mathbf{x}\| \leqslant \|\mathbf{y}\|$). This latter condition ensures that $\|\mathbf{c}^+\|$ is a convex function (see Question 14.2) and includes all L_p and scaled L_p norms for $1 \leqslant p \leqslant \infty$. Thus the results of this section are not just restricted to polyhedral norms. It is convenient to introduce the concept of a *dual norm*

$$\|\mathbf{u}\|_D \triangleq \max_{\|\mathbf{v}\| \leqslant 1} \mathbf{u}^{\mathrm{T}}\mathbf{v}. \tag{14.3.6}$$

The dual of $\|\cdot\|_1$ is $\|\cdot\|_\infty$, and vice versa, and the $\|\cdot\|_2$ is self-dual. Expressions for the subdifferential of the functions $\|\mathbf{c}\|$ and $\|\mathbf{c}^+\|$ are given by

$$\partial\|\mathbf{c}\| = \{\boldsymbol{\lambda}: \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{c} = \|\mathbf{c}\|, \|\boldsymbol{\lambda}\|_D \leqslant 1\} \tag{14.3.7}$$

and

$$\partial\|\mathbf{c}^+\| = \{\boldsymbol{\lambda}: \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{c} = \|\mathbf{c}^+\|,\ \boldsymbol{\lambda} \geqslant \mathbf{0},\ \|\boldsymbol{\lambda}\|_D \leqslant 1\}. \tag{14.3.8}$$

The proof of these expressions is sketched out in some detail in Questions 14.4 and 14.5. Expressions (14.1.12) to (14.1.15) are special cases of (14.3.7) and (14.3.8). The main results giving the equivalence between local solutions of (14.3.1) and (14.3.2) (or between (14.3.3) and (14.3.4)) can now be given. In fact the latter result only is given; the former case is similar (but easier). A point to bear in mind is that in these theorems $\boldsymbol{\lambda}^*$ refers to a multiplier vector for the constrained problem (14.3.3) and $v\boldsymbol{\lambda}^*$ is the equivalent multiplier vector for the exact penalty function problem (14.3.4).

Theorem 14.3.1

If $v < 1/\|\boldsymbol{\lambda}^*\|_D$ and $\mathbf{c}^{*+} = \mathbf{0}$ then the second order sufficient conditions at $\mathbf{x}^*$ for problems (14.3.3) *and* (14.3.4) *are equivalent. Therefore if they hold, the fact that* $\mathbf{x}^*$ *solves* (14.3.3) *implies that* $\mathbf{x}^*$ *solves* (14.3.4), *and vice versa.*

Proof

Second order conditions are given for problem (14.3.3) in Theorem 9.3.2 and for problem (14.3.4) in Theorem 14.2.3. The first order requirements are that $\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}$, $\boldsymbol{\lambda}^* = \mathbf{0}$, $\boldsymbol{\lambda}^{*\mathrm{T}}\mathbf{c}^* = \|\mathbf{c}^{*+}\| = 0$ (with a suitable sign change), and by $v\mathbf{g}^* + \mathbf{A}^*(v\boldsymbol{\lambda}^*) = \mathbf{0}$, $v\boldsymbol{\lambda}^* \in \partial h^*$ respectively, which are clearly equivalent from (14.3.8) if $v < 1/\|\boldsymbol{\lambda}^*\|_D$. Next the sets G^* defined in (9.3.11) (with $\|\mathbf{s}\|_2 = 1$) and (14.2.20) are shown to be equivalent. For convenience refer to these as G_9^* and G_{14}^* respectively. Let $\mathbf{s} \in G_{14}^*$. From (14.2.20) and the above it follows that

$$\mathbf{s}^{\mathrm{T}}\mathbf{A}^*(\boldsymbol{\lambda} - v\boldsymbol{\lambda}^*) \leqslant 0 \qquad \forall\, \boldsymbol{\lambda} \in \partial\|\mathbf{c}^{*+}\|. \tag{14.3.9}$$

Let $\mathscr{A}^*$ denote active constraints at $\mathbf{x}^*$ in (14.3.3). For $i \in \mathscr{A}^*$ and small ε it follows using $\|v\boldsymbol{\lambda}^*\|_D < 1$ that $\boldsymbol{\lambda} = v\boldsymbol{\lambda}^* + \varepsilon\mathbf{e}_i \in \partial\|\mathbf{c}^{*+}\|$ either if $\lambda_i^* = 0$ and $\varepsilon > 0$ or if $\lambda_i^* > 0$ and $\pm\varepsilon > 0$. Hence from (14.3.9),

$$\begin{aligned} \mathbf{s}^{\mathrm{T}}\mathbf{a}_i^* &\leqslant 0 \quad \text{if } \lambda_i^* = 0, \\ \mathbf{s}^{\mathrm{T}}\mathbf{a}_i^* &= 0 \quad \text{if } \lambda_i^* > 0, \end{aligned} \qquad i \in \mathscr{A}^*. \tag{14.3.10}$$

Thus $\mathbf{s} \in G_9^*$. Conversely let $\mathbf{s} \in G_9^*$. It follows from (14.3.10) that $\mathbf{s}^{\mathrm{T}}\mathbf{a}_i^*\lambda_i^* = 0$ which also implies that $\mathbf{s}^{\mathrm{T}}\mathbf{g}^* = 0$ from the first order conditions. Let $\boldsymbol{\lambda} \in \partial h^*$; then $\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{c}^* = 0$ and $\boldsymbol{\lambda} \geqslant \mathbf{0}$ imply that if $c_i^* < 0$ then $\lambda_i = 0$ so constraints $i \notin \mathscr{A}^*$ can be ignored. Otherwise $\mathbf{s}^{\mathrm{T}}\mathbf{a}_i^*\lambda_i \leqslant 0$ follows from (14.3.10) and hence $\max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) = 0$. Thus $\mathbf{s} \in G_{14}^*$ and the equivalence of G_9^* and G_{14}^* is shown. Finally conditions (9.3.15) and (14.2.24) are clearly equivalent so the equivalence of the second order conditions is demonstrated. $\square$

The proof of this theorem is very similar to that of Theorem 14.2.4, and indeed the results are very similar, showing as they do the equivalence between a composite NSO problem and a nonlinear programming problem. Similar theorems relating the solutions of (14.3.3) and (14.3.4) are given by Charalambous (1979) and Han and Mangasarian (1979).

The requirement that second order conditions hold and that $\nu < 1/\|\boldsymbol{\lambda}^*\|_D$ in Theorem 14.3.1 cannot easily be relaxed, as the following simple examples show. In all of these, $x^* = 0$ solves (14.3.3) but does not solve (14.3.4) using $\|\cdot\|_1$ for the reasons given. For the problem: min x subject to $x^2 \leqslant 0$, x^* is not a KT point. For the problem: min x^3 subject to $x^5 \geqslant 0$, x^* is a KT point and $\lambda^* = 0$ but the curvature condition (9.3.14) is not strict. For the problem: min $x - \frac{1}{2}x^2$ subject to $0 \leqslant x \leqslant 1$, x^* is a KT point with $\lambda^* = 1$. Then if $\nu = 1$ the condition $\nu < 1/\|\boldsymbol{\lambda}^*\|_D$ is not strict for L_1 norm. The last example illustrates another fact. The proof of Theorem 14.3.1 shows that $G_9^* \subset G_{14}^*$ without requiring that $\nu < 1/\|\boldsymbol{\lambda}^*\|_D$. (In the last example G_9^* is empty whilst G_{14}^* is the direction $s = -1$.) Thus if $\mathbf{x}^*$ satisfies second order sufficient conditions for (14.3.4) it follows that both $\nu \leqslant 1/\|\boldsymbol{\lambda}^*\|_D$ and $\mathbf{x}^*$ satisfies second order sufficient conditions for (14.3.3). Finally if $\nu > 1/\|\boldsymbol{\lambda}^*\|_D$ then a result going the other way can be proved.

Theorem 14.3.2

If first order conditions $\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}$ *hold for* (14.3.3) *and if* $\nu > 1/\|\boldsymbol{\lambda}^*\|_D$ *then* $\mathbf{x}^*$ *is not a local minimizer of* (14.3.4).

Proof

Since $\|\nu\boldsymbol{\lambda}^*\|_D > 1$, the vector $\nu\boldsymbol{\lambda}^*$ is not in $\partial\|\mathbf{c}^{*+}\|$ and so the vector $\nu\mathbf{g}^* + \mathbf{A}^*(\nu\boldsymbol{\lambda}^*) = \mathbf{0}$ is not in $\partial\phi^*$. Hence by (14.2.6) $\mathbf{x}^*$ is not a local minimizer of $\phi(\mathbf{x})$. □

Another important result concerns the regularity assumption in (14.2.21). If $\mathbf{x}^*$, $\boldsymbol{\lambda}^*$ satisfy first order conditions for either (14.3.1) or (14.3.3) (including feasibility) then under mild assumptions (14.2.21) holds, even though the norm may not be polyhedral. Again the result is proved only in the more difficult case.

Lemma 14.3.1

If $\mathbf{x}^*, \boldsymbol{\lambda}^*$ *is a KT point for* (14.3.3), *if the vectors* $\mathbf{a}_i^*$, $i \in \mathscr{A}^*$, *are linearly independent, and if* $\|\nu\boldsymbol{\lambda}^*\|_D < 1$, *then* $\mathscr{G}^* = G^*$.

Proof

Let $\mathbf{s} \in G^*$. Then as in the proof of Theorem 14.3.1, (14.3.10) holds. By the independence assumption there exists an arc $\mathbf{x}(\theta)$ with $\mathbf{x}(0) = \mathbf{x}^*$ and $\dot{\mathbf{x}}(0) = \mathbf{s}$, for which

$$\left.\begin{array}{ll} c_i(\mathbf{x}(\theta)) = 0 & \text{if } \lambda_i^* > 0, \\ c_i(\mathbf{x}(\theta)) \leqslant 0 & \text{if } \lambda_i^* = 0, \end{array}\right. \quad i \in \mathscr{A}^*$$

for $\theta \geqslant 0$ and sufficiently small (see the proof of Lemma 9.2.2). It follows that $\| \mathbf{c}(\mathbf{x}(\theta))^+ \| = \mathbf{c}(\mathbf{x}(\theta))^{\mathrm{T}} \boldsymbol{\lambda}^* = 0$ so $\mathbf{x}(\theta)$ is feasible in (14.2.19) and hence $\mathbf{s} \in \mathscr{G}^*$.

An alternative and more general proof of this lemma can also be given. If $h(\mathbf{c}) = \|\mathbf{c}\|$ and $\mathbf{c}^* = \mathbf{0}$ then $\partial h^* = \{\boldsymbol{\lambda} : \|\boldsymbol{\lambda}\|_D \leqslant 1\}$ and hence a basis for $\partial h^* - \boldsymbol{\lambda}^*$ is the set of columns $\mathbf{e}_j$ $j = 1, 2, \ldots, m$. If $h(\mathbf{c}) = \|\mathbf{c}^+\|$ and $\mathbf{c}^{*+} = \mathbf{0}$ then from (14.3.8) it follows that $\partial h^* = \{\boldsymbol{\lambda} : c_i^* < 0 \Rightarrow \lambda_i = 0, \boldsymbol{\lambda} \geqslant 0, \|\boldsymbol{\lambda}\|_D \leqslant 1\}$. In both cases therefore a basis for $\partial h^* - \boldsymbol{\lambda}^*$ is the set of vectors $\mathbf{e}_i$ $i \in \mathscr{A}^*$. Thus independence of the vectors $\mathbf{a}_i^*$ $i \in \mathscr{A}^*$ implies that the rank condition (14.2.31) holds. The positive homogeneity of the functions $\|\cdot\|$ or $\|\cdot^+\|$ implies that $h(\mathbf{c})$ is locally linear about $\mathbf{c}^*$ when $\mathbf{c}^* = \mathbf{0}$ or $\mathbf{c}^{*+} = \mathbf{0}$. Also the KT conditions and $\|\nu\boldsymbol{\lambda}^*\|_D \leqslant 1$ imply that $\mathbf{x}^*$ satisfies first order conditions (14.2.18). Thus the assumptions of Lemma 14.2.7 are valid and it therefore follows that $\mathscr{G}^* = G^*$. □

To summarize all these results, if $\mathbf{x}^*$ is a feasible point of the nonlinear programming problem and if $\nu > 0$ is sufficiently small, then except in limiting cases, $\mathbf{x}^*$ solves the nonlinear programming problem if and only if $\mathbf{x}^*$ minimizes the corresponding exact penalty function (locally speaking).

In practice there are considerable advantages in using the L_1 norm to define (14.3.2) or (14.3.4). More details of this particular case, including an alternative derivation of the optimality conditions without using subgradients, is given in Section 12.3.

14.4 ALGORITHMS

This section reviews progress in the development of algorithms for many different classes of NSO problem. It happens that similar ideas have been tried in different situations and it is convenient to discuss them in a unified way. At present there is a considerable amount of interest in these developments and it is not possible to say yet what the best approaches are. This is particularly true when curvature estimates are updated by quasi-Newton like schemes. However some important common features are emerging which this review tries to bring out.

Many methods are line search methods in which on each iteration a direction of search $\mathbf{s}^{(k)}$ is determined from some model situation, and $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{s}^{(k)}$ is obtained by choosing $\alpha^{(k)}$ to minimize approximately the objective function $\phi(\mathbf{x}^{(k)} + \alpha\mathbf{s}^{(k)})$ along the line (see Section 2.3). A typical line search algorithm (various have been suggested) operates under similar principles to those described in Section 2.6 and uses a combination of sectioning and interpolation, but there are some new features. One is that since $\phi(\mathbf{x})$ may contain a polyhedral component $h(\mathbf{c}(\mathbf{x}))$, the interpolating function must also have the same type of structure. The simplest possibility is to interpolate a one variable function of the form

$$\psi(\alpha) = q(\alpha) + h(\mathbf{l}(\alpha)) \tag{14.4.1}$$

where $q(\alpha)$ is quadratic and the functions $\mathbf{l}(\alpha)$ are linear. For composite NSO applications q and $\mathbf{l}$ can be estimated from information about f and $\mathbf{c}$, and $\alpha^{(k)}$ is then determined by minimizing (14.4.1). For basic NSO, only the values of ϕ and $d\phi/d\alpha$ are known at any point so it must be assumed that $\psi(\alpha)$ has a more simple structure, for example the max of just two linear functions $l_i(\alpha)$. Many other possibilities exist. Another new feature concerns what acceptability tests to use, analogous to (2.5.1), (2.5.2), (2.5.4), and (2.5.6) for smooth functions. If the line search minimum is non-smooth (see Figure 14.4.1) then it is not appropriate to try to make the directional derivative small, as (2.5.6) does, since such a point may not exist. Many line searches use the Wolfe–Powell conditions (2.5.1) and (2.5.4). For this choice the range of acceptable α-values is the interval $[a, b]$ in Figure 14.4.1, and this should be compared with Figure 2.5.2. It can be seen that this line search has the effect that the acceptable point *always overshoots the minimizing value of* α, and in fact this may occur by a substantial amount, so as to considerably slow down the rate of convergence of the algorithm in which it is embedded. I have found it valuable to use a different test (Fletcher, 1981) that the line search is terminated when the predicted reduction based on (14.4.1) is sufficiently small, and in particular when the predicted reduction on any subinterval is no greater than 0.1 times the total reduction in $\phi(\mathbf{x})$ so far achieved in the search. This condition has been found to work well in ensuring that $\alpha^{(k)}$ is close to the minimizing value of α.

It is now possible to examine algorithms for NSO problems in many variables. In basic NSO only a limited amount of information is available at any given point $\mathbf{x}$, namely $\phi(\mathbf{x})$ and one element $\mathbf{g} \in \partial\phi(\mathbf{x})$ (usually $\nabla\phi(\mathbf{x})$ since $\phi(\mathbf{x})$ is almost everywhere differentiable). This makes the basic NSO problem more difficult than composite NSO in which values of f and $\mathbf{c}$ (and their derivatives)

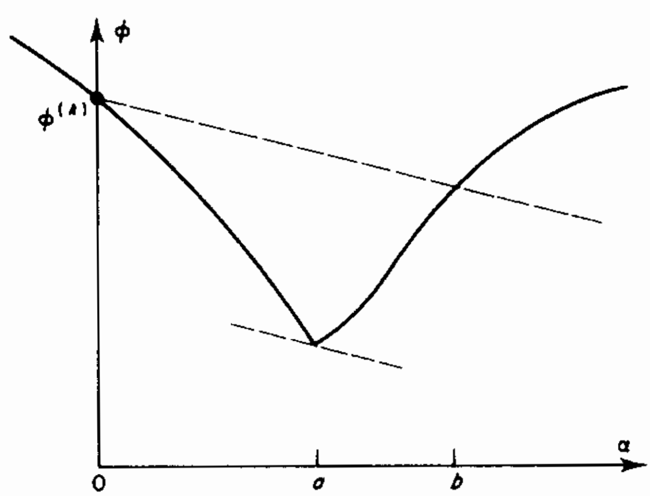

Figure 14.4.1 Line search for non-smooth functions

are given in (14.1.1). The simplest method for basic NSO is an analogue of the steepest descent method in which $\mathbf{s}^{(k)} = -\mathbf{g}^{(k)}$ is used as a search direction; this is known as *subgradient optimization*. Because of its simplicity the method has received much attention both in theory and in practice (see the references given in Lemarechal and Mifflin (1978)), but it is at best linearly convergent, as illustrated by the results for smooth functions (see Section 2.3). In fact the situation is worse in the non-smooth case because the convergence result (Theorem 2.4.1) no longer holds. In fact examples of non-convergence are easily constructed. Assume that the line search is exact and that the subgradient obtained at $\mathbf{x}^{(k+1)}$ corresponds to the piece which is active for $\alpha \geqslant \alpha^{(k)}$. Then the example

$$\begin{aligned}
\phi(\mathbf{x}) &= \max_{i=1,2,3} c_i(\mathbf{x}) \\
c_1(\mathbf{x}) &= -5x_1 + x_2 \\
c_2(\mathbf{x}) &= x_1^2 + x_2^2 + 4x_2 \\
c_3(\mathbf{x}) &= 5x_1 + x_2
\end{aligned} \tag{14.4.2}$$

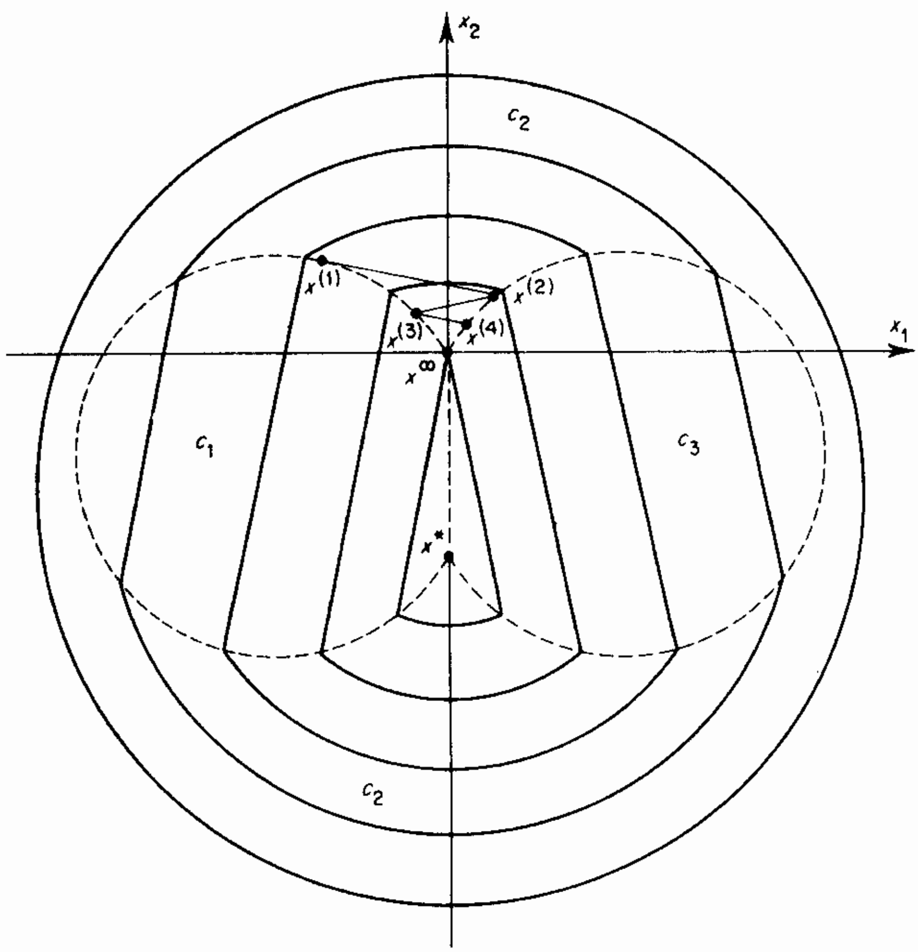

Figure 14.4.2 False convergence of steepest-descent-like methods for problem (14.4.2)

due to Demyanov and Malozemov (1971) illustrates convergence from the initial point $\mathbf{x}^{(1)}$ (see Figure 14.4.2) to the non-optimal point $\mathbf{x}^{\infty} = \mathbf{0}$. The solution is at $\mathbf{x}^* = (0, -3)^{\mathrm{T}}$. At $\mathbf{x}^{(1)}, \mathbf{x}^{(3)}, \ldots$ the subgradient corresponding to c_1 is used and at $\mathbf{x}^{(2)}, \mathbf{x}^{(4)}, \ldots$ the subgradient corresponding to c_3. It is easily seen from Figure 14.4.2 that the sequence $\{\mathbf{x}^{(k)}\}$ oscillates between the two curved surfaces of non-differentiability and converges to $\mathbf{x}^{\infty}$ which is not optimal. In fact it is not necessary for the surfaces of non-differentiability to be curved, and a similar polyhedral (piecewise linear) example can easily be constructed for which the algorithm also fails.

It can be argued that a closer analogue to the steepest descent method at a point of non-differentiability is to search along the the steepest descent direction $\mathbf{s}^{(k)} = -\bar{\mathbf{g}}^{(k)}$ where $\bar{\mathbf{g}}^{(k)}$ minimizes $\|\mathbf{g}\|_2$ for $\mathbf{g} \in \partial\phi^{(k)}$. This interpretation is justified for convex functions in (14.2.12) and a similar result holds when the composite function $\phi(\mathbf{x})$ is in use. Let $\partial\phi^{(k)}$ be defined by the convex hull of its extreme points, $\partial\phi^{(k)} = \operatorname{conv} \mathbf{g}_i^{(k)}$, say. (For composite functions $\mathbf{g}_i^{(k)} = \mathbf{g}^{(k)} + \mathbf{A}^{(k)}\mathbf{h}_i$, $i \in \mathscr{A}^{(k)}$, see (14.2.16).) Then $\bar{\mathbf{g}}^{(k)}$ is defined by solving the problem

$$\begin{aligned} &\underset{\mathbf{g},\boldsymbol{\mu}}{\text{minimize}} && \mathbf{g}^{\mathrm{T}}\mathbf{g} \\ &\text{subject to} && \mathbf{g} = \sum_i \mu_i \mathbf{g}_i^{(k)}, \\ &&& \sum_i \mu_i = 1, \qquad \boldsymbol{\mu} \geq \mathbf{0}, \end{aligned} \tag{14.4.3}$$

and $\mathbf{s}^{(k)} = -\bar{\mathbf{g}}^{(k)}$. This problem is similar to a least distance QP problem and is readily solved by the methods of Section 10.5. The resulting method terminates finitely when $\phi(\mathbf{x})$ is polyhedral (piecewise linear) and exact line searches are used. This latter condition ensures that the $\mathbf{x}^{(k)}$ are on the surfaces of non-differentiability for which $\partial\phi^{(k)}$ has more than one element. The algorithm is idealized however in that exact line searches are not generally possible in practice. Also the whole subdifferential $\partial\phi^{(k)}$ is not usually available, and even if it were the non-exact line search would cause $\partial\phi^{(k)} = \nabla\phi^{(k)}$ to hold and the method would revert to subgradient optimization.

The spirit of this type of method however is preserved in *bundle* methods (Lemarechal, 1978). A bundle method is a line search method which solves subproblem (14.4.3) to define $\mathbf{s}^{(k)}$, except that the vectors $\mathbf{g}_i^{(k)}$ are elements of a 'bundle' B rather than the extreme points of $\partial\phi^{(k)}$. Initially B is set to $B = \mathbf{g}^{(1)} \in \partial\phi^{(1)}$, and in the simplest form of the algorithm, subgradients $\mathbf{g}^{(2)}$, $\mathbf{g}^{(3)}, \ldots$ are added to B on successive iterations. The method continues in this way until $\mathbf{0} \in B$. Then the bundle B is reset, for instance to the current $\mathbf{g}^{(k)}$, and the iteration is continued. With careful manipulation of B (see for example the *conjugate subgradient method* of Wolfe (1975)) a convergence result for the algorithm can be proved and a suitable termination test obtained. However Wolfe's method does not terminate finitely at the solution for polyhedral $\phi(\mathbf{x})$. An example (due to Powell) has the pentagonal contours illustrated in Figure 14.4.3. At $\mathbf{x}^{(1)}$, B is $\mathbf{g}_1$, and at $\mathbf{x}^{(2)}$, B is $\{\mathbf{g}_1, \mathbf{g}_3\}$. At $\mathbf{x}^{(3)}$, B is $\{\mathbf{g}_1, \mathbf{g}_3, \mathbf{g}_4\}$ and $\mathbf{0} \in B$ so the process is restarted with $B = \mathbf{g}_3$ (or $B = \mathbf{g}_4$ with the same

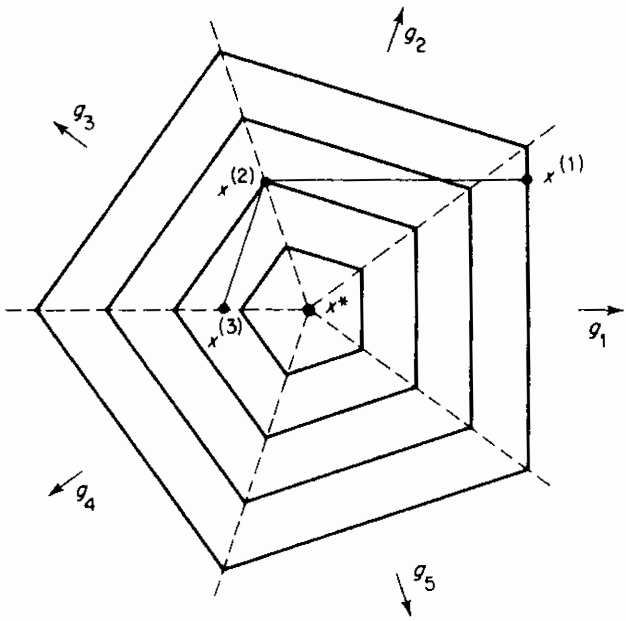

Figure 14.4.3 Non-termination of a bundle method

conclusions). The whole sequence is repeated and the points $\mathbf{x}^{(k)}$ cycle round the pentagon, approaching the centre but without ever terminating there.

This non-termination can be corrected by including more elements in B when restarting (for example both $\mathbf{g}_3$ and $\mathbf{g}_4$) or alternatively deleting old elements like $\mathbf{g}_1$. This can be regarded as providing a better estimate of $\partial\phi^{(k)}$ by B. However, it is by no means certain that this would give a better method: in fact the idealized steepest descent method itself can fail to converge. Problem (14.4.2) again illustrates this; this time the search directions are tangential to the surfaces of non-differentiability and oscillatory convergence to $\mathbf{x}^\infty$ occurs. This example illustrates the need to have information about the subdifferential at $\partial\phi^\infty$ (which is lacking in this example) to avoid the false convergence at $\mathbf{x}^\infty$. Generalizations of bundle methods have therefore been suggested which attempt to enlarge $\partial\phi^{(k)}$ so as to contain subgradient information at neighbouring points. One possibility is to use the ε-subdifferential.

$$\partial_\varepsilon\phi(\mathbf{x}) = \{\mathbf{g}\colon \phi(\mathbf{x}+\boldsymbol{\delta}) \geqslant \phi(\mathbf{x}) + \mathbf{g}^{\mathrm{T}}\boldsymbol{\delta} - \varepsilon \quad \forall\, \boldsymbol{\delta}\}$$

which contains $\partial\phi(\mathbf{x})$. Bundle methods in which B approximates this set are discussed by Lemarechal (1978) and are currently being researched.

A quite different type of method for basic NSO is to use the information $\mathbf{x}^{(k)}$, $\phi^{(k)}$, $\mathbf{g}^{(k)}$ obtained at any point to define the linear function

$$\phi^{(k)} + (\mathbf{x}-\mathbf{x}^{(k)})^{\mathrm{T}}\mathbf{g}^{(k)} \tag{14.4.4}$$

and to use these linearizations in an attempt to model $\phi(\mathbf{x})$. If $\phi(\mathbf{x})$ is convex then this function is a supporting hyperplane and this fact is exploited in the *cutting plane method*. The linearizations are used to define a model polyhedral convex function and the minimizer of this function determines the next iterate.

Specifically the linear program

$$\begin{aligned}&\underset{\mathbf{x},v}{\text{minimize}} \quad v\\ &\text{subject to} \quad v \geqslant \phi^{(i)} + (\mathbf{x}-\mathbf{x}^{(i)})^{\mathrm{T}}\mathbf{g}^{(i)}, \qquad i=1,2,\ldots,k\end{aligned} \tag{14.4.5}$$

is solved to determine $\mathbf{x}^{(k+1)}$. Then the linearization determined by $\mathbf{x}^{(k+1)}$, $\phi^{(k+1)}$, and $\mathbf{g}^{(k+1)}$ is added to the set and the process is repeated. On the early iterations a step restriction $\|\mathbf{x}-\mathbf{x}^{(k)}\|_\infty \leqslant h$ is required to ensure that (14.4.5) is not unbounded. A line search can also be added to the method. A similar method can be used on non-convex problems if old linearizations are deleted in a systematic way.

More sophisticated algorithms for basic NSO are discussed at the end of this section and attention is now turned to algorithms for minimizing the composite function (14.1.1) (which includes max functions, etc., when $h(\mathbf{c})$ is the polyhedral function (14.1.2)—see the special cases in (14.1.3)). Linear approximations have also been used widely and the simplest method is to replace $\mathbf{c}(\mathbf{x})$ by the first order Taylor series approximation

$$\mathbf{c}(\mathbf{x}^{(k)}+\boldsymbol{\delta}) \approx \mathbf{l}^{(k)}(\boldsymbol{\delta}) = \mathbf{c}^{(k)} + \mathbf{A}^{(k)\mathrm{T}}\boldsymbol{\delta} \tag{14.4.6}$$

and $f(\mathbf{x})$ (if present) by

$$f(\mathbf{x}^{(k)}+\boldsymbol{\delta}) \approx f^{(k)} + \mathbf{g}^{(k)\mathrm{T}}\boldsymbol{\delta}, \tag{14.4.7}$$

and to substitute these approximations into (14.1.6) which is an equivalent form of the original problem. The linear program

$$\begin{aligned}&\underset{\boldsymbol{\delta},v}{\text{minimize}} \quad v\\ &\text{subject to} \quad v - (\mathbf{g}^{(k)} + \mathbf{A}^{(k)}\mathbf{h}_i)^{\mathrm{T}}\boldsymbol{\delta} \geqslant f^{(k)} + \mathbf{c}^{(k)\mathrm{T}}\mathbf{h}_i + b_i \qquad \forall i\end{aligned} \tag{14.4.8}$$

is solved to determine $\boldsymbol{\delta}^{(k)}$ and hence $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}$. The only difference between this and (14.4.5) is that here there is sufficient information available at $\mathbf{x}^{(k)}$ to determine linear approximations to all the pieces, whereas in basic NSO this information has to be accumulated on a sequence of iterations. When applied to $\min \|\mathbf{c}(\mathbf{x})\|_\infty$ or $\min \|\mathbf{c}(\mathbf{x})\|_1$ the iteration based on (14.4.8) can be considered as a *Gauss–Newton method* by analogy with the same method for solving $\min \|\mathbf{c}(\mathbf{x})\|_2^2$ (see Section 6.1) which is based on the same linear approximations. An early study of this type of method is by Osborne and Watson (1969) and a more elaborate recent method is given by Charalambous and Conn (1978). As with Gauss–Newton methods, convergence is not guaranteed but this can likewise be rectified by going to a restricted step type of method. In this case $\|\mathbf{x}-\mathbf{x}^{(k)}\|_\infty \leqslant h^{(k)}$ is a suitable choice since it preserves a linear programming subproblem. This approach has been investigated by Madsen (1975).

Linear approximation methods are most successful when the linearizations of the active pieces at $\mathbf{x}^*$ fully determine $\mathbf{x}^*$ (that is to say when the first order conditions are sufficient, which occurs when the set G^* in (14.2.20) corresponding to directions of zero slope is empty and curvature effects are negligible). A necessary condition for this is that there are $n+1$ (or more) active pieces at $\mathbf{x}^*$. In these circumstances the order of convergence of the method based on (14.4.8) is second order (to prove this, show that the method is ultimately equivalent to the Newton–Raphson method for solving the active constraints in (14.4.8) or the structure functional constraints in (14.2.35)). This situation is most likely to occur in problems such as over-determined L_1 or L_∞ data fitting, where these methods can be very successful.

In general however it is not possible to exclude curvature effects, and when these are significant then methods based only on first order information converge slowly and become unreliable. It is therefore important to consider how second order information can be used to obtain second order convergence. Now it has been shown in Theorem 14.2.4 that the composite NSO problem is equivalent to the nonlinear programming problem (14.2.35). It is also known from Theorem 12.4.1 that the SQP method converges at second order for a nonlinear programming problem. Thus by writing down an NSO subproblem which is locally equivalent to the SQP subproblem, a method which converges at second order for the composite NSO problem can be given. The quadratic function

$$q^{(k)}(\boldsymbol{\delta}) = \tfrac{1}{2}\boldsymbol{\delta}^{\mathrm{T}}\mathbf{W}^{(k)}\boldsymbol{\delta} + \mathbf{g}^{(k)\mathrm{T}}\boldsymbol{\delta} + f^{(k)} \tag{14.4.9}$$

is again needed, and the SQP method for (14.2.35) with iterates $(\mathbf{x}^{(k)}, \mathbf{u}^{(k)})$ would involve the subproblem

$$\begin{aligned} &\underset{\boldsymbol{\delta}}{\text{minimize}} \quad q^{(k)}(\boldsymbol{\delta}) + \boldsymbol{\delta}^{\mathrm{T}}\mathbf{A}^{(k)}\boldsymbol{\lambda}^* \\ &\text{subject to} \quad \mathbf{D}^{*\mathrm{T}}(\mathbf{c}^* - \mathbf{l}^{(k)}(\boldsymbol{\delta})) = \mathbf{0}. \end{aligned} \tag{14.4.10}$$

(Here the result is used that $\mathbf{W}^{(k)} = \nabla^2 f_i^{(k)} + \sum \lambda_i^{(k)} \nabla^2 c_i^{(k)} = \nabla^2 f_i^{(k)} + \sum \lambda_i^* \nabla^2 c_i^{(k)} + \sum u_j^{(k)} \nabla^2 (\mathbf{D}^{*\mathrm{T}}\mathbf{c})_j^{(k)}$ where $\boldsymbol{\lambda}^{(k)} = \boldsymbol{\lambda}^* + \mathbf{D}^*\mathbf{u}^{(k)}$.) It is shown below under mild assumptions that this is equivalent to iterating with $(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)})$ using the NSO subproblem

$$\underset{\boldsymbol{\delta}}{\text{minimize}} \quad \psi^{(k)}(\boldsymbol{\delta}) \triangleq q^{(k)}(\boldsymbol{\delta}) + h(\mathbf{l}^{(k)}(\boldsymbol{\delta})) \tag{14.4.11}$$

The kth iteration is simply to find the minimizer, $\boldsymbol{\delta}^{(k)}$ say, of (14.4.11) and set $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}$, whilst $\boldsymbol{\lambda}^{(k+1)}$ is set as the multipliers of this subproblem. This method is referred to here as the *sequential non-smooth quadratic programming (SNQP) method.* The first reference I have to such a method is that it is mentioned briefly by Pshenichnyi (1978) in the context of minimizing max functions, but it was first given in this general form in the previous edition of this book, where it was referred to as the *QL method.* The $\mathrm{SL_1QP}$ method in Section 12.4 is another example of such a method applied to minimizing non-smooth L_1 problems. The SNQP method can be seen to have a nice interpretation. The function $\psi^{(k)}(\boldsymbol{\delta})$ in (14.4.11) approximates the composite function $\phi(\mathbf{x})$ local to $\mathbf{x}^{(k)}$. This approximation is constructed by replacing $\mathbf{c}(\mathbf{x})$ in (14.1.1) or (14.2.34) by its linear approximation $\mathbf{l}^{(k)}(\boldsymbol{\delta})$ and $f(\mathbf{x})$ by the quadratic approximation $q^{(k)}(\boldsymbol{\delta})$. This quadratic has the Hessian matrix $\mathbf{W}^{(k)}$ which accounts for curvature

Table 14.4.1 Application of the SNQP method to problem (12.3.2)

k	1	2	3	4	5	6
$\mathbf{x}^{(k)}$	2	1.25	0.804310	0.712691	0.707133	0.707107
	0	0.5	0.801724	0.712981	0.707127	0.707107
$\lambda^{(k)}$	1	0.625	0.622844	0.692599	0.706968	0.707107
$\phi^{(k)}$	1	-0.9375	-1.316358	-1.409402	-1.414194	-1.414214
$\boldsymbol{\delta}^{(k)}$	-0.75	-0.445690	-0.091619	-0.005558	-0.000026	$2_{10}-10$
	0.5	0.301724	-0.088744	-0.005854	-0.000020	$-9_{10}-10$

in the functions $f(\mathbf{x})$ and $c_i(\mathbf{x})$. This is exactly the same way as the SQP method handles the equivalent nonlinear programming problem (14.2.35). Moreover in Section 12.4 the SQP method is itself shown to be equivalent to the Newton–Raphson method applied to the first order conditions that arise in the method of Lagrange multipliers. Thus the SNQP method can be regarded as a generalization of the Newton–Raphson method that is appropriate for composite NSO problems.

An illustration of the SNQP method applied to solve problem (12.3.2) with $v = 1$ is given. Initial approximations $\mathbf{x}^{(1)} = (2, 0)^T$ and $\lambda^{(1)} = 1$ are taken and the progress of the iterations is given in Table 14.4.1. It can be seen that the convergence is second order with $\|\boldsymbol{\delta}^{(k+1)}\|_2 / \|\boldsymbol{\delta}^{(k)}\|_2^2 \approx 0.5$. The contours of the approximating function $\psi^{(k)}$ for $k = 1$ and $k = 2$ are illustrated in Figure 14.4.4. Each function has quadratic pieces with a discontinuous derivative on a linear surface (partly dotted) which is the linearization (through (14.4.6)) of the unit circle on which the discontinuity in Figure 12.3.1 occurs.

It is possible to under mild conditions to prove that the SNQP method conveges locally at second order in all cases of any real interest. Two preliminary lemmas are required and the main result then follows.

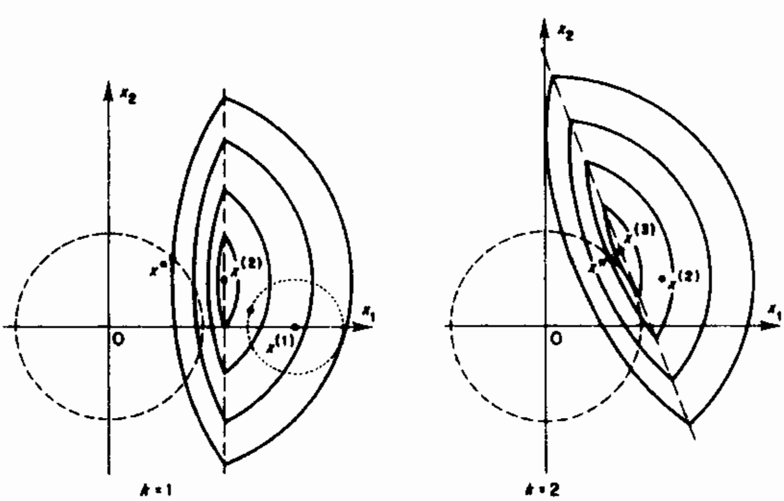

Figure 14.4.4 Contours of the approximating function in the SNQP method

Lemma 14.4.1

If $h(\mathbf{c})$ is locally linear about $\mathbf{c}^*$ and if $\mathbf{c}'\in\Omega$ then

$$\partial h(\mathbf{c}') \subseteq \partial h(\mathbf{c}^*). \tag{14.4.12}$$

Proof

Define $\mathbf{c}_\theta = \mathbf{c}^* + \theta(\mathbf{c}' - \mathbf{c}^*)$ where $\theta \geqslant 0$ and $\mathbf{c}_\theta \in \Omega$. Using the locally linear property (14.2.29)

$$\begin{aligned} h(\mathbf{c}_\theta) &= h(\mathbf{c}^*) + \max_{\boldsymbol{\lambda}\in\partial h^*} \theta(\mathbf{c}' - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda} \\ &= h(\mathbf{c}^*) + \theta \max_{\boldsymbol{\lambda}\in\partial h^*} (\mathbf{c}' - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda} \\ &= h(\mathbf{c}^*) + \theta(h(\mathbf{c}') - h(\mathbf{c}^*)) \end{aligned} \tag{14.4.13}$$

Consider $\boldsymbol{\lambda}\in\partial h(\mathbf{c}')$. The subgradient inequality gives

$$h(\mathbf{c}) \geqslant h(\mathbf{c}') + (\mathbf{c} - \mathbf{c}')^{\mathrm{T}}\boldsymbol{\lambda} \qquad \forall\, \mathbf{c} \tag{14.4.14}$$

so

$$h(\mathbf{c}_\theta) \geqslant h(\mathbf{c}') + (\mathbf{c}_\theta - \mathbf{c}')^{\mathrm{T}}\boldsymbol{\lambda}.$$

Substituting for $\mathbf{c}_\theta$ and using (14.4.13) gives

$$(1-\theta)(h(\mathbf{c}^*) - h(\mathbf{c}')) \geqslant (1-\theta)(\mathbf{c}^* - \mathbf{c}')^{\mathrm{T}}\boldsymbol{\lambda}.$$

Since Ω is open, $\exists$ points $\mathbf{c}_\theta\in\Omega$ with both $\theta < 1$ and $\theta > 1$, which implies that

$$h(\mathbf{c}^*) - h(\mathbf{c}') = (\mathbf{c}^* - \mathbf{c}')^{\mathrm{T}}\boldsymbol{\lambda}. \tag{14.4.15}$$

It follows from (14.4.14) that

$$h(\mathbf{c}) \geqslant h(\mathbf{c}^*) + (\mathbf{c} - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda} \qquad \forall\, \mathbf{c} \tag{14.4.16}$$

and hence $\boldsymbol{\lambda}\in\partial h(\mathbf{c}^*)$. □

Lemma 14.4.2

If $h(\mathbf{c})$ is locally linear at $\mathbf{c}^*$, if $\mathbf{c}'\in\Omega$, and if

$$\mathbf{D}^{*\mathrm{T}}(\mathbf{c}^* - \mathbf{c}') = \mathbf{0} \tag{14.4.17}$$

then

$$\partial h(\mathbf{c}') = \partial h(\mathbf{c}^*). \tag{14.4.18}$$

Proof

Let $\boldsymbol{\lambda}'$ denote an arbitrary vector in $\partial h(\mathbf{c}')$, so that $\boldsymbol{\lambda}'\in\partial h(\mathbf{c}^*)$ and

$$h(\mathbf{c}^*) - h(\mathbf{c}') = (\mathbf{c}^* - \mathbf{c}')^{\mathrm{T}}\boldsymbol{\lambda}' \tag{14.4.19}$$

both hold, from Lemma 14.4.1 and (14.4.15). In view of Lemma 14.4.1, only the reverse inclusion needs to be proved, so consider $\boldsymbol{\lambda} \in \partial h(\mathbf{c}^*)$. From (14.2.30) $\exists$ a vector $\mathbf{v}$ such that $\boldsymbol{\lambda} - \boldsymbol{\lambda}' = \mathbf{D}^*\mathbf{v}$ and hence from (14.4.17)

$$\mathbf{c}^{*\mathrm{T}}(\boldsymbol{\lambda} - \boldsymbol{\lambda}') = \mathbf{c}'^{\mathrm{T}}(\boldsymbol{\lambda} - \boldsymbol{\lambda}'). \tag{14.4.20}$$

The subgradient inequality (14.4.16) holds and it follows from (14.4.19) and (14.4.20) that (14.4.14) holds, and hence $\boldsymbol{\lambda} \in \partial h(\mathbf{c}')$. $\square$

Theorem 14.4.1 (Second order convergence of the SNQP method)

Let $\mathbf{x}^*, \boldsymbol{\lambda}^*$ *satisfy the second order sufficient conditions of Theorem* 14.2.3, *and assume that* $h(\mathbf{c})$ *is locally linear at* $\mathbf{c}^*$, *and that both the rank assumption* (14.2.31) *and strict complementarity* $\mathbf{0} \in \operatorname{int} U^*$ *hold. In a neighbourhood of* $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ *the SNQP subproblem* (14.4.11) *has a locally unique minimizer. Assuming that this point is located when solving the subproblem, then the SNQP method is equivalent to the SQP method applied to* (14.2.35) *based on the subproblem* (14.4.10). *The iterates in the SNQP method converge to* $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ *and the order of convergence is quadratic.*

Proof

Consider the system of equations

$$\begin{bmatrix} \mathbf{W} & \mathbf{AD}^* \\ (\mathbf{AD}^*)^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \boldsymbol{\delta} \\ \mathbf{u}^+ \end{pmatrix} = \begin{pmatrix} -\mathbf{g} - \mathbf{A}\boldsymbol{\lambda}^* \\ \mathbf{D}^{*\mathrm{T}}(\mathbf{c}^* - \mathbf{c}) \end{pmatrix} \tag{14.4.21}$$

where $\mathbf{c} = \mathbf{c}(\mathbf{x})$, $\mathbf{A} = \mathbf{A}(\mathbf{x})$, etc. and $\mathbf{W} = \mathbf{W}(\mathbf{x}, \boldsymbol{\lambda})$, and where $(\mathbf{x}, \boldsymbol{\lambda})$ is in some neighbourhood of $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$. From second order sufficient conditions and strict complementarity it follows that

$$\mathbf{s}^{\mathrm{T}}\mathbf{W}^*\mathbf{s} > 0 \qquad \forall \mathbf{s} \in \mathcal{N}^*, \tag{14.4.22}$$

and hence by continuity that

$$\mathbf{s}^{\mathrm{T}}\mathbf{W}\mathbf{s} > 0 \qquad \forall \mathbf{s} \in \mathcal{N} = \{\mathbf{s}\colon \|\mathbf{s}\|_2 = 1,\ \mathbf{s}^{\mathrm{T}}\mathbf{AD}^* = \mathbf{0}^{\mathrm{T}}\} \tag{14.4.23}$$

(see Question 12.17 for example). It follows from (14.4.22) and the rank assumption (14.2.31) that the matrix

$$\begin{bmatrix} \mathbf{W}^* & \mathbf{A}^*\mathbf{D}^* \\ (\mathbf{A}^*\mathbf{D}^*)^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \tag{14.4.24}$$

is non-singular (see Question 12.4). Hence by the implicit function theorem there exists a neighbourhood of $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ such that for any $(\mathbf{x}, \boldsymbol{\lambda})$ in this neighbourhood, there is a uniquely defined continuous solution $(\boldsymbol{\delta}, \mathbf{u}^+)$ to (14.4.21), and at $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ the solution satisfies $\boldsymbol{\delta} = \mathbf{0}$ and $\mathbf{u}^+ = \mathbf{0}$. Equations (14.4.21) are the first order conditions for $(\boldsymbol{\delta}, \mathbf{u}^+)$ to solve the SQP subproblem (14.4.10) as defined for an iterate $(\mathbf{x}, \boldsymbol{\lambda})$. If $\boldsymbol{\lambda}^+ = \boldsymbol{\lambda}^* + \mathbf{D}^*\mathbf{u}^+$, the first equation in (14.4.21) is also the Lagrangian first order condition for $(\boldsymbol{\delta}, \boldsymbol{\lambda}^+)$ in the SNQP subproblem (14.4.11) as defined for an iterate $(\mathbf{x}, \boldsymbol{\lambda})$. Because $\mathbf{0} \in \operatorname{int} U^*$ it follows by continuity that $\mathbf{u}^+ \in U^*$ and hence $\boldsymbol{\lambda}^+ \in \partial h(\mathbf{c}^*)$. From the second condition in (14.4.21), continuity of $\boldsymbol{\delta}$ and Lemma 14.4.2, it follows that $\boldsymbol{\lambda}^+ \in \partial h(\mathbf{c} + \mathbf{A}^T\boldsymbol{\delta})$, and hence $(\boldsymbol{\delta}, \boldsymbol{\lambda}^+)$ satisfies first order conditions for $(\boldsymbol{\delta}, \boldsymbol{\lambda}^+)$ to solve the SNQP problem. Second order sufficient conditions for both problems involve (14.4.23) which is known to hold. It follows that a unique local minimizer exists for both subproblems, and in fact it is easily seen to be a global minimizer for the SQP subproblem. Thus if $\boldsymbol{\lambda} = \boldsymbol{\lambda}^* + \mathbf{D}^*\mathbf{u}$ and $\boldsymbol{\lambda}^+ = \boldsymbol{\lambda}^* + \mathbf{D}^*\mathbf{u}^+$, then calculating $(\boldsymbol{\delta}, \mathbf{u}^+)$ from an iterate $(\mathbf{x}, \mathbf{u})$ by solving the SQP subproblem is equivalent to calculating $(\boldsymbol{\delta}, \boldsymbol{\lambda}^+)$ from an iterate $(\mathbf{x}, \boldsymbol{\lambda})$ by solving the SNQP subproblem, assuming that the locally unique solution is located when solving the latter (see Question 14.16). Finally consider applying the SNQP method from an iterate $(\mathbf{x}^{(1)}, \boldsymbol{\lambda}^{(1)})$ sufficiently close to $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$. Then as above, we can express $\boldsymbol{\lambda}^{(2)} = \boldsymbol{\lambda}^* + \mathbf{D}^*\mathbf{u}^{(2)}$. It then follows that for $k \geqslant 2$ the iterates $(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)})$ in the SNQP method are equivalent to the iterates $(\mathbf{x}^{(k)}, \mathbf{u}^{(k)})$ in the SQP method, assuming that $\boldsymbol{\lambda}^{(k)} = \boldsymbol{\lambda}^* + \mathbf{D}^*\mathbf{u}^{(k)}$. Convergence and second order convergence follow from Theorem 12.4.1 $\square$

The assumptions in Theorem 14.4.1 are mostly mild ones, the only significant restriction being that $h(\mathbf{c})$ must be locally linear at $\mathbf{c}^*$. As has been pointed out in Sections 14.2 and 14.3, this includes not only the polyhedral convex functions of Sections 14.1 but also the exact penalty function problems for smooth norms in Section 14.3. This covers all convex composite problems that are of any real interest. A final remark about order of convergence is that when $h(\mathbf{c}) = \|\mathbf{c}\|$ is a smooth norm (for example $\|\cdot\|_2$) and $\mathbf{c}^* \neq \mathbf{0}$ then $\mathscr{G}^* \neq G^*$ and the above development is not appropriate. In this case the most appropriate second order sufficient conditions are those for a smooth problem that $\nabla\phi^* = \mathbf{0}$ and $\nabla^2\phi^*$ is positive definite. However the SNQP method does not reduce to Newton's method so it is an open question as to whether second order convergence can be deduced.

These results on second order convergence make it clear that the SNQP method is attractive as a starting point for developing a general purpose algorithm for composite NSO problems. However the basic method itself is not robust and can fail to converge (when $m = 0$ it is just the basic Newton's method—see Section 3.1). An obvious precaution to prevent divergence is to require that the sequence $\{\phi^{(k)}\}$ is non-increasing. One possibility is to use the correction from solving (14.4.11) as a direction of search $\mathbf{s}^{(k)}$ along which to approximately minimize the objective function $\phi(\mathbf{x}^{(k)} + \alpha\mathbf{s}^{(k)})$. Both the basic method and the line search modification can fail in another way however, in that remote from the solution the function $\psi^{(k)}(\boldsymbol{\delta})$ may not have a minimizer. This is analogous to the case that $\mathbf{G}^{(k)}$ is indefinite in smooth unconstrained minimization. An effective modification in the latter case is the restricted step or trust region approach (Chapter 5) and it is fruitful to consider modifying the SNQP method in this way. Since (14.4.11) contains no side conditions the incorporation of a step restriction causes no difficulty. This is illustrated in Figure 14.4.4 ($k = 1$) in which the dotted circle with centre $\mathbf{x}^{(1)}$ is a possible trust region. Clearly there is no difficulty in minimizing $\psi^{(1)}(\boldsymbol{\delta})$ within this region: the solution is the point on the periphery of the circle. Contrast this situation with Section 12.4 where a trust region cannot be used in conjunction with the SQP method because there may not exist feasible points in the resulting problem. Fletcher (1982a) describes a prototype restricted step method for solving composite NSO problems which is globally convergent without the need to assume that vectors $\mathbf{a}_i^\infty$ are independent or multipliers $\boldsymbol{\lambda}^{(k)}$ are bounded. This algorithm is described in more detail in Section 14.5 where the global convergence property is justified. Good practical experience with the method in L_1 exact penalty function applications is reported by Fletcher (1981), including the solution of test problems on which other methods fail.

There are two main computational costs with the SNQP method. One is that associated with solving the subproblem (14.4.11) and this of course depends on the form of the composite function $h(\cdot)$. When h is a polyhedral convex function the subproblem is equivalent to a QP problem and so the complexity of the subproblem is roughly the same as that for QP. In the L_1 case the subproblem becomes the L_1QP problem described in more detail in Section 12.3. The other main cost is the need to evaluate the problem functions f and $\mathbf{c}$ and their first and second derivatives at each iteration. The need to provide particularly second derivatives may be a disadvantage, but there is no difficulty in approximating the second derivative matrix $\mathbf{W}^{(k)}$ for instance by finite differences or by using a quasi-Newton update. One possibility is to use the modified BFGS formula given by Powell (1978a) as described in Section 12.4, and this is quite straightforward. Likewise the possibility of approximating the *reduced* second derivative matrix also suggests itself but it is by no means clear how to use this matrix to the best advantage in a general SNQP context. However, Womersley (1981) gives a method applicable to the particular case of max function.

There is one adverse feature of the SNQP method, and hence of any other superlinearly convergent algorithm for composite NSO which is not present with similar algorithms for smooth optimization. In the latter case when $\mathbf{x}^{(k)}$ is close to $\mathbf{x}^*$, the unit step of Newton's method reduces the objective function, so no line search or trust region restriction is brought into play to force the objective function to decrease. Thus the second order convergence, which depends on taking the unit step, is unaffected. The same situation does *not* hold for composite NSO as Maratos (1978) observes (see also Mayne (1980), Chamberlain *et al.* (1982) and Question 14.13). In particular there are well behaved composite NSO problems in which $(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)})$ can be arbitrarily close to $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ and the unit step of the SNQP method can fail to reduce $\phi(\mathbf{x})$. The *Maratos effect* as it might be called, thus causes the unit step to be rejected and can obviate the possibility of obtaining second order convergence. More precisely for an iteration of the SNQP method, define the *actual reduction*

$$\Delta\phi^{(k)} = \phi(\mathbf{x}^{(k)}) - \phi(\mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}), \tag{14.4.25}$$

the *predicted reduction*

$$\Delta\psi^{(k)} = \psi^{(k)}(\mathbf{0}) - \psi^{(k)}(\boldsymbol{\delta}^{(k)}) = \phi(\mathbf{x}^{(k)}) - \psi^{(k)}(\boldsymbol{\delta}^{(k)}) \tag{14.4.26}$$

and the ratio

$$r^{(k)} = \Delta\phi^{(k)}/\Delta\psi^{(k)}. \tag{14.4.27}$$

Also let

$$\varepsilon^{(k)} = \max(\|\mathbf{x}^{(k)} - \mathbf{x}^*\|, \|\boldsymbol{\lambda}^{(k)} - \boldsymbol{\lambda}^*\|). \tag{14.4.28}$$

Then for smooth unconstrained optimization (e.g. Theorems 3.1.1 and 5.2.2) it is possible to get both $\varepsilon^{(k+1)} = O(\varepsilon^{(k)^2})$ and $r^{(k)} = 1 + O(\varepsilon^{(k)})$, whereas for composite NSO these results are inconsistent, and indeed $\Delta\phi^{(k)} < 0$ and hence $r^{(k)} < 0$ can occur which corresponds to the Maratos, effect. The reason for this is illustrated in Figure 14.4.5 in the non-smooth case. If $\mathbf{x}^{(k)}$ is close to (or on) a groove, the predicted first order changes to ϕ are small (or zero). Thus the predicted second order changes to ϕ can be significant. However, the functions $c_i(\mathbf{x})$ are only approximated linearly in calculating $\mathbf{x}^{(k+1)}$ and hence second order errors arise in $\phi(\mathbf{x}^{(k+1)})$ which can dominate the second order changes in the predicted reduction. This effect is due to the presence of derivative discontinuities in $\phi(\mathbf{x})$. These discontinuities also can occasionally cause algorithms for SNQP to converge slowly (for example see Table 12.4.3 for $\mathrm{SL}_1\mathrm{QP}$). Both effects are most likely to be observed when the second order errors are large, that is when the grooves are significantly nonlinear, such as for example when the penalty term in an exact penalty function is relatively large. Thus if robust algorithms for NSO are to be developed, it is important to consider how these effects can be overcome.

The above observations indicate that a modification which corrects the second order errors in c_i is appropriate. This can be done (Fletcher, 1982b) by making a single 'second order correction' (SOC) step, which corrects the 'basic step' obtained by solving (14.4.11). Similar ideas occur in Coleman and Conn (1982a, b) and indirectly in Chamberlain *et al.* (1982). The idea can be simply explained as follows. The basic step $\boldsymbol{\delta}$ in the SNQP method is obtained by solving the

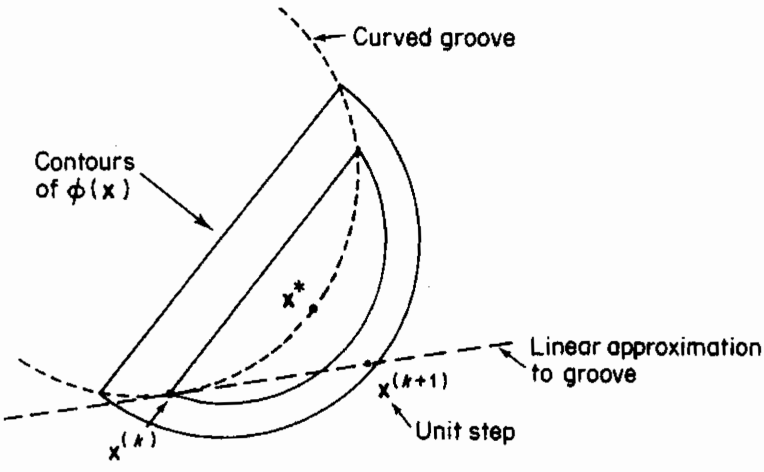

Figure 14.4.5 Maratos effect

subproblem

$$\underset{\delta}{\text{minimize}} \quad q^{(k)}(\delta) + h([c_i^{(k)} + \mathbf{a}_i^{(k)\mathrm{T}}\delta]). \tag{14.4.31}$$

where $[\cdot]$ denotes the vector having the given elements. To account satisfactorily for second order changes it would be preferable to solve

$$\underset{\delta}{\text{minimize}} \quad q^{(k)}(\delta) + h([c_i^{(k)} + \mathbf{a}_i^{(k)\mathrm{T}}\delta + \tfrac{1}{2}\delta^{\mathrm{T}}\mathbf{G}_i^{(k)}\delta]) \tag{14.4.32}$$

where $\mathbf{G}_i = \nabla^2 c_i$, but this subproblem is intractable. However, if the solution to (14.4.31) is $\hat{\delta}$, and we define $\hat{\mathbf{x}} = \mathbf{x}^{(k)} + \hat{\delta}$ and evaluate $\hat{\mathbf{c}} = \mathbf{c}(\hat{\mathbf{x}})$, then to second order

$$\hat{c}_i = c_i^{(k)} + \mathbf{a}_i^{(k)\mathrm{T}}\hat{\delta} + \tfrac{1}{2}\hat{\delta}^{\mathrm{T}}\mathbf{G}_i^{(k)}\hat{\delta}. \tag{14.4.33}$$

Equation (14.4.33) can be rearranged to provide an expression for $\frac{1}{2}\hat{\delta}^{\mathrm{T}}\mathbf{G}_i^{(k)}\hat{\delta}$ which can be used to estimate the second order term in (14.4.32). Thus the SOC subproblem

$$\underset{\delta}{\text{minimize}} \quad q^{(k)}(\delta) + h([\hat{c}_i - \mathbf{a}_i^{(k)\mathrm{T}}\hat{\delta} + \mathbf{a}_i^{(k)\mathrm{T}}\delta]) \tag{14.4.34}$$

is suggested, and the step $\delta^{(k)}$ that is used in the resulting algorithm to calculate $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta^{(k)}$ is that obtained by solving (14.4.34). Note that the only difference between (14.4.34) and (14.4.31) is a change to the constant term in the argument of $h(\cdot)$. For polyhedral problems this merely has the effect of shifting the linear approximation of the curved groove so that $\mathbf{x}^{(k+1)}$ is closer to the groove, as illustrated in Figure 14.4.6. This figure also shows that the SOC step is related to the projection step computed in feasible direction methods for nonlinear programming (cf. Figure 12.5.1) and in particular to that given by (12.5.11) for $\alpha^{(k)} = 1$.

The SOC modification has a number of practical advantages. It uses the

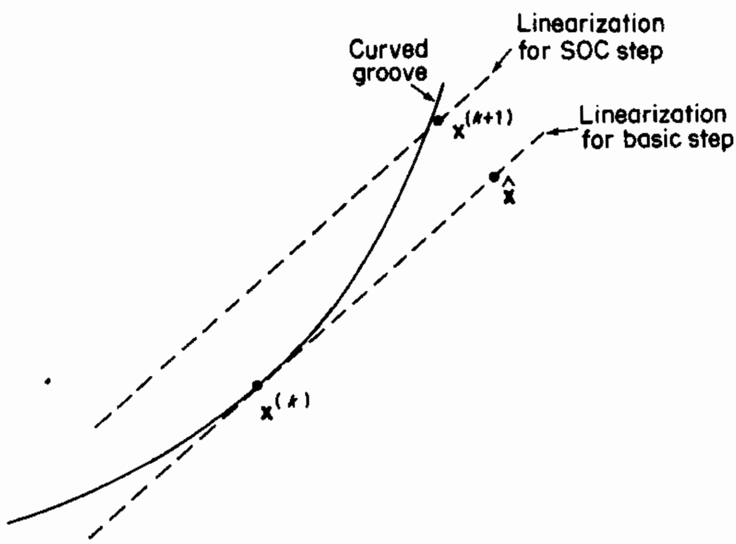

Figure 14.4.6 The second order correction step

same software for calculating both the basic and SOC step and so is easy to code. No extra derivative matrices need be calculated to define the SOC step and only one additional evaluation of the active constraint residuals $\mathbf{c}(\hat{\mathbf{x}})$ is required. Moreover, and of great importance, only a change to the constant values $\mathbf{c}^{(k)}$ in the basic step subproblem is made. Assuming that the software for the subproblem enables parametric changes to be made efficiently, it is likely that the same active set and same matrix factorizations can be used to solve the SOC subproblem directly. In these circumstances the SOC step can be computed at negligible cost. The above method represents the simplest way to use an SOC step: Fletcher (1982b) also suggests a more complicated algorithm in which the basic step is accepted for $\boldsymbol{\delta}^{(k)}$ in certain circumstances, and the SOC step is only used if it is judged to be necessary. An SOC step can also be used effectively when the SQP method is used in conjunction with a non-smooth penalty function.

Other second order methods for NSO have also been suggested, in particular the use of a *hybrid method* (or *2-stage method*). Stage 1 is the Gauss–Newton trust region method (14.4.8 ff.) which converges globally so is potentially reliable, but often converges only at first order which can be slow. Stage 2 is Newton's method applied to the first order conditions. This converges locally at second order but does not have any global properties. A hybrid method is one which switches between these methods and aims to combine their best features. To apply Newton's method requires a knowledge of the active pieces in the NSO problem (essentially a set of structure functionals as in (14.2.35) is required) and this knowledge can also be gained from the progress of the Gauss–Newton iteration. Hybrid methods are given by Watson (1979) for the minimax problem and by McLean and Watson (1980) for the L_1 problem. Hybrid methods using a quasi-Newton method for stage 2 are given by Hald and Madsen (1981, 1985). A recent method for general non-smooth problems, using the Coleman–Conn method for stage 2, is given by Fletcher and Sainz de la Maza (1987). Hybrid methods can work well but I anticipate one disadvantage. If the NSO problem is such that neighbourhood of the solution in which Newton's method works well is small, then most of the time will be taken up in stage 1, using the first order method. If this converges slowly then the hybrid method will not solve the problem effectively. Thus it is important to ensure that the stage 2 method is used to maximum effect.

The above techniques are only appropriate for composite NSO problems, and it is also important to consider applications to basic NSO in which second order information has been used. Womersley (1978) describes a method which requires $\phi^{(i)}$, $\nabla\phi^{(i)}$, and $\nabla^2\phi^{(i)}$ to be available for any $\mathbf{x}^{(i)}$, $i = 1, 2, \ldots, k$, and each such set of information is assumed to arise from an active piece of $\phi(\mathbf{x})$. This information is used to build up a linear approximation for each piece valid at $\mathbf{x}^{(k)}$, and also to give the matrix $\mathbf{W}^{(k)}$. This is then used as in (14.4.10). Information about pieces may be rejected if the Lagrange multipliers $\mu_i^{(k+1)}$ determine that the piece is no longer active. Also a Levenberg–Marquardt modification (Section 5.2) is made to $\mathbf{W}^{(k)}$ to ensure positive definiteness. A difficulty with the algorithm is that repeated approximations to the same piece tend to be collected, which cause degeneracy in the QP solver close to the solution. A modification is proposed in which an extra item of information is supplied at each $\mathbf{x}^{(i)}$ which is an integer *label* of the active piece. Such information can be supplied in most basic NSO applications. This labelling enables a single most recent approximation to each piece to be maintained and circumvents the degeneracy problem. Again this approach is considerably more effective than the subgradient or cutting plane methods as numerical evidence given by Womersley suggests (~ 20 iterations as against ~ 300 on a typical problem). A quasi-Newton version of the algorithm is described in Womersley (1981).

14.5 A GLOBALLY CONVERGENT PROTOTYPE ALGORITHM

The aim of this section is to show that the SNQP method based on solving the subproblem (14.4.11) can be incorporated readily with the idea of a step restriction which is known to give good numerical results in smooth unconstrained optimization (Chapter 5). Subproblem (14.4.11) contains curvature information and potentially allows second order convergence to occur, whereas the step length restriction is shown to ensure global convergence. The resulting method (Fletcher, 1981, 1982a) is applicable to the solution of all composite NSO problems, including exact penalty functions, best L_1 and L_∞ approximation, etc., but excluding basic NSO problems. The term 'prototype algorithm' indicates that the algorithm is presented in a simple format as a means of making its convergence properties clear. It admits of the algorithm being modified to improve its practical performance whilst not detracting from these properties. The motivation for using a step restriction is that it defines a *trust region* for the correction $\boldsymbol{\delta}$ by

$$\|\boldsymbol{\delta}\| \leqslant h^{(k)} \tag{14.5.1}$$

in which the Taylor series approximations (14.4.6) and (14.4.9) are assumed to be adequate. The norm in (14.5.1) is arbitrary but either the $\|\cdot\|_\infty$ or the $\|\cdot\|_2$ is the most likely choice, especially the former since (14.5.2) below can then be solved by QP-like methods (see (10.3.5) and Section 12.3). On each iteration the subproblem is to minimize (14.4.11) subject to this restriction, that is

$$\begin{aligned} &\underset{\boldsymbol{\delta}}{\text{minimize}} \quad \psi^{(k)}(\boldsymbol{\delta}) \\ &\text{subject to} \quad \|\boldsymbol{\delta}\| \leqslant h^{(k)}. \end{aligned} \tag{14.5.2}$$

The radius $h^{(k)}$ of the trust region is adjusted adaptively to be as large as possible subject to adequate agreement between $\phi(\mathbf{x}^{(k)} + \boldsymbol{\delta})$ and $\psi^{(k)}(\boldsymbol{\delta})$ being maintained. This can be quantified by defining the *actual reduction*

$$\Delta\phi^{(k)} = \phi(\mathbf{x}^{(k)}) - \phi(\mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}) \tag{14.5.3}$$

and the *predicted reduction*

$$\Delta\psi^{(k)} = \phi(\mathbf{x}^{(k)}) - \psi^{(k)}(\boldsymbol{\delta}^{(k)}). \tag{14.5.4}$$

Then the ratio

$$r^{(k)} = \frac{\Delta\phi^{(k)}}{\Delta\psi^{(k)}} \tag{14.5.5}$$

measures the extent to which ϕ and $\psi^{(k)}$ agree local to $\mathbf{x}^{(k)}$. The rules for changing $h^{(k)}$ in the model algorithm are those given in Section 5.1 and are not elaborated on further, except to emphasize that in practice the rule for reducing $h^{(k)}$ can be more elaborate, based perhaps on some sort of interpolation (Fletcher 1981).

The kth iteration of the prototype algorithm is as follows.

(i) Given $\mathbf{x}^{(k)}$, $\boldsymbol{\lambda}^{(k)}$, and $h^{(k)}$, calculate $f^{(k)}$, $\mathbf{g}^{(k)}$, $\mathbf{c}^{(k)}$, $\mathbf{A}^{(k)}$, and $\mathbf{W}^{(k)}$ which determine $\phi^{(k)}$ and $\psi^{(k)}(\boldsymbol{\delta})$.
(ii) Find a global solution $\boldsymbol{\delta}^{(k)}$ to (14.5.2).
(iii) Evaluate $\phi(\mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)})$ and calculate $\Delta\phi^{(k)}$, $\Delta\psi^{(k)}$, and $r^{(k)}$.
(iv) If $r^{(k)} < 0.25$ set $h^{(k+1)} = \|\boldsymbol{\delta}^{(k)}\|/4$,
if $r^{(k)} > 0.75$ and $\|\boldsymbol{\delta}^{(k)}\| = h^{(k)}$ set $h^{(k+1)} = 2h^{(k)}$, (14.5.6)
otherwise set $h^{(k+1)} = h^{(k)}$.
(v) If $r^{(k)} \leqslant 0$ set $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$, $\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)}$,
else $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \boldsymbol{\delta}^{(k)}$, $\boldsymbol{\lambda}^{(k+1)} =$ multipliers from (14.5.2).

The parameters 0.25, 0.75, etc., which arise are arbitrary and are not very sensitive but the values given here are typical. The solution of constrained NSO problems like (14.5.2) is considered in more detail in Section 14.6 but it is a straightforward extension of the unconstrained case and multipliers $\boldsymbol{\lambda}^{(k+1)} \in \partial h(\mathbf{l}^{(k)}(\boldsymbol{\delta}^{(k)}))$ exist at the solution in an analogous way to Theorem 14.2.1. It is these multipliers that are used in step (v).

In proving global convergence, a result is used relating the directional derivative (14.2.14) of the composite function (14.2.13) and the difference quotient between two points $\mathbf{x}^{(k)}$ and $\mathbf{x}^{(k)} + \varepsilon^{(k)}\mathbf{s}$ in a common direction $\mathbf{s}$ ($\neq \mathbf{0}$), as both points approach a fixed point $\mathbf{x}'$. The result is a special case of one due to Clarke (1975) but is proved here directly for completeness.

Lemma 14.5.1

Let S be the set of all sequences $\mathbf{x}^{(k)} \to \mathbf{x}'$, $\varepsilon^{(k)} \to 0$, and let $\mathbf{c}^{(k)} \triangleq \mathbf{c}(\mathbf{x}^{(k)})$, etc., and $\mathbf{c}_\varepsilon^{(k)} \triangleq \mathbf{c}(\mathbf{x}^{(k)} + \varepsilon^{(k)}\mathbf{s})$; then

$$\limsup_{S} \frac{h(\mathbf{c}_\varepsilon^{(k)}) - h(\mathbf{c}^{(k)})}{\varepsilon^{(k)}} = \max_{\boldsymbol{\lambda}\in\partial h'} \mathbf{s}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\lambda} \tag{14.5.7}$$

in the sense that the difference quotient is bounded above and the sup of all accumulation points of the quotient taken over all sequences in S is the directional derivative $\max_{\boldsymbol{\lambda}\in\partial h'} \mathbf{s}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\lambda}$.

Proof

By the integral form of the Taylor series

$$\mathbf{c}_\varepsilon^{(k)} = \mathbf{c}^{(k)} + \varepsilon^{(k)} \int_0^1 [\mathbf{A}(\mathbf{x}^{(k)} + \theta\varepsilon^{(k)}\mathbf{s})]^{\mathrm{T}}\mathbf{s}\,\mathrm{d}\theta = \mathbf{c}^{(k)} + \varepsilon^{(k)}\mathbf{d}^{(k)} \tag{14.5.8}$$

say, where $\mathbf{d}^{(k)} \to \mathbf{A}'^{\mathrm{T}}\mathbf{s}$ as $k \to \infty$. Let $\boldsymbol{\lambda}_\varepsilon^{(k)} \in \partial h_\varepsilon^{(k)} \triangleq \partial h(\mathbf{c}_\varepsilon^{(k)})$. Using the subgradient inequality and (14.5.8),

$$h(\mathbf{c}^{(k)}) \geqslant h(\mathbf{c}_\varepsilon^{(k)}) + (\mathbf{c}^{(k)} - \mathbf{c}_\varepsilon^{(k)})^{\mathrm{T}}\boldsymbol{\lambda}_\varepsilon^{(k)} = h(\mathbf{c}_\varepsilon^{(k)}) - \varepsilon^{(k)}\mathbf{d}^{(k)\mathrm{T}}\boldsymbol{\lambda}_\varepsilon^{(k)}$$

or

$$\mathbf{d}^{(k)\mathrm{T}}\boldsymbol{\lambda}_\varepsilon^{(k)} \geqslant \frac{h(\mathbf{c}_\varepsilon^{(k)}) - h(\mathbf{c}^{(k)})}{\varepsilon^{(k)}}. \tag{14.5.9}$$

Since $\partial h_\varepsilon^{(k)}$ is bounded in a neighbourhood of $\mathbf{x}'$ (Lemma 14.2.1) the difference quotient is bounded above. Now consider a subsequence for which the difference quotient accumulates, and let

$$\lim \frac{h(\mathbf{c}_\varepsilon^{(k)}) - h(\mathbf{c}^{(k)})}{\varepsilon^{(k)}} > \max_{\boldsymbol{\lambda} \in \partial h'} \mathbf{s}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\lambda}. \tag{14.5.10}$$

Since $\boldsymbol{\lambda}_\varepsilon^{(k)}$ is bounded there exists a thinner subsequence for which $\boldsymbol{\lambda}_\varepsilon^{(k)} \to \boldsymbol{\lambda}'$ and since $\mathbf{c}_\varepsilon^{(k)} \to \mathbf{c}'$ it follows by Lemma 14.2.4 that $\boldsymbol{\lambda}' \in \partial h'$. Thus from (14.5.9),

$$\lim \frac{h(\mathbf{c}_\varepsilon^{(k)}) - h(\mathbf{c}^{(k)})}{\varepsilon^{(k)}} \leqslant \mathbf{s}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\lambda}' \leqslant \max_{\boldsymbol{\lambda} \in \partial h'} \mathbf{s}^{\mathrm{T}}\mathbf{A}'\boldsymbol{\lambda} \tag{14.5.11}$$

which contradicts (14.5.10), so that the reverse inequality ($\leqslant$) in (14.5.10) is true. Finally by taking $\mathbf{x}^{(k)} = \mathbf{x}'$ and $\varepsilon^{(k)} \downarrow 0$, the discussion before (14.2.14) shows that there is a sequence in S which attains equality with the directional derivative. Thus (14.5.7) is established. $\square$

Corollary

Define $\phi(\mathbf{x})$ *by* (14.2.13) *where* $f \in \mathbb{C}^1$. *Then*

$$\limsup_{S} \frac{\phi(\mathbf{x}^{(k)} + \varepsilon^{(k)}\mathbf{s}) - \phi(\mathbf{x}^{(k)})}{\varepsilon^{(k)}} = \max_{\boldsymbol{\lambda} \in \partial h'} \mathbf{s}^{\mathrm{T}}(\mathbf{g}' + \mathbf{A}'\boldsymbol{\lambda}). \tag{14.5.12}$$

Proof

The result follows by using an analogous Taylor series for $f(\mathbf{x})$ as in the proof of the lemma. $\square$

It is now possible to state the main result of this section.

Theorem 14.5.1 (Global convergence of prototype algorithm)

Let $\mathbf{x}^{(k)} \in B \subset \mathbb{R}^n \forall k$ *where* B *is bounded and let* f, $\mathbf{c}$ *be* $\mathbb{C}^2$ *functions whose second derivative matrices are bounded on* B. *Then there exists an accumulation point* $\mathbf{x}^\infty$ *of algorithm* (14.5.6) *at which the first order conditions of Theorem* 14.2.1 *hold, that is equivalently*

$$\max_{\boldsymbol{\lambda} \in \partial h^\infty} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^\infty + \mathbf{A}^\infty\boldsymbol{\lambda}) \geqslant 0 \quad \forall \mathbf{s}. \tag{14.5.13}$$

Proof

There exists a convergent subsequence $\mathbf{x}^{(k)} \to \mathbf{x}^\infty$ for which either

(i) $r^{(k)} < 0.25, h^{(k+1)} \to 0$, and hence $\|\boldsymbol{\delta}^{(k)}\| \to 0$ or
(ii) $r^{(k)} \geqslant 0.25$ and $\inf h^{(k)} > 0$.

In either case (14.5.13) is shown to hold. In case (i) let $\exists$ a descent direction $\mathbf{s}$ ($\|\mathbf{s}\| = 1$) at $\mathbf{x}^\infty$, that is

$$\max_{\boldsymbol{\lambda} \in \partial h^\infty} \mathbf{s}^T(\mathbf{g}^\infty + \mathbf{A}^\infty \boldsymbol{\lambda}) = -d, \qquad d > 0. \tag{14.5.14}$$

By Taylor series

$$\begin{aligned} f(\mathbf{x}^{(k)} + \varepsilon^{(k)}\mathbf{s}) &= f^{(k)} + \varepsilon^{(k)}\mathbf{s}^T\mathbf{g}^{(k)} + o(\varepsilon^{(k)}) \\ &= q^{(k)}(\varepsilon^{(k)}\mathbf{s}) + o(\varepsilon^{(k)}) \end{aligned} \tag{14.5.15}$$

by (14.4.9), since $\boldsymbol{\lambda}^{(k)}$ is bounded by Lemma 14.2.1 and $\nabla^2 f$, $\nabla^2 c_i$ are bounded by assumption. Likewise by (14.4.6),

$$\mathbf{c}(\mathbf{x}^{(k)} + \varepsilon^{(k)}\mathbf{s}) = \mathbf{l}^{(k)}(\varepsilon^{(k)}\mathbf{s}) + o(\varepsilon^{(k)}) \tag{14.5.16}$$

and hence by (14.1.2), the boundedness of ∂h, and (14.4.11), it follows that

$$\begin{aligned} \phi(\mathbf{x}^{(k)} + \varepsilon^{(k)}\mathbf{s}) &= q^{(k)}(\varepsilon^{(k)}\mathbf{s}) + h(\mathbf{l}^{(k)}(\varepsilon^{(k)}\mathbf{s}) + o(\varepsilon^{(k)})) + o(\varepsilon^{(k)}) \\ &= q^{(k)}(\varepsilon^{(k)}\mathbf{s}) + h(\mathbf{l}^{(k)}(\varepsilon^{(k)}\mathbf{s})) + o(\varepsilon^{(k)}) \\ &= \psi^{(k)}(\varepsilon^{(k)}\mathbf{s}) + o(\varepsilon^{(k)}). \end{aligned} \tag{14.5.17}$$

Writing $\varepsilon^{(k)} = \|\boldsymbol{\delta}^{(k)}\|$ and considering a step along $\mathbf{s}$ in the subproblem, it follows by the optimality of $\boldsymbol{\delta}^{(k)}$ that

$$\begin{aligned} \Delta\psi^{(k)} &\geqslant \phi^{(k)} - \psi^{(k)}(\varepsilon^{(k)}\mathbf{s}) \\ &= \phi^{(k)} - \phi(\mathbf{x}^{(k)} + \varepsilon^{(k)}\mathbf{s}) + o(\varepsilon^{(k)}) \\ &\geqslant \varepsilon^{(k)}(d + o(1)) + o(\varepsilon^{(k)}) = d\varepsilon^{(k)} + o(\varepsilon^{(k)}) \end{aligned} \tag{14.5.18}$$

by the corollary to Lemma 14.5.1 and (14.5.14). But (14.5.17) implies that

$$\Delta\phi^{(k)} = \Delta\psi^{(k)} + o(\varepsilon^{(k)})$$

and hence $r^{(k)} = \Delta\phi^{(k)}/\Delta\psi^{(k)} = 1 + o(\varepsilon^{(k)})/\Delta\psi^{(k)} = 1 + o(1)$ from (14.5.18) since $d > 0$, which contradicts $r^{(k)} < 0.25$. Thus $d \leqslant 0$ for all $\mathbf{s}$ and hence (14.5.13) holds at $\mathbf{x}^\infty$.

In case (ii) the argument of case (ii) of Theorem 5.1.1 is largely followed (see also Fletcher, 1982a). The only extra step is to deduce from $\mathbf{x}^{(k)} \in B$ that $\partial h^{(k)}$ is uniformly bounded for all k, so that the parameters $\boldsymbol{\lambda}^{(k)}$ are bounded. Thus a thinner subsequence can be chosen such that $\boldsymbol{\lambda}^{(k)} \to \boldsymbol{\lambda}^\infty$ and hence $\mathbf{W}^{(k)} \to \mathbf{W}^\infty$. Then functions $q^\infty(\boldsymbol{\delta})$, $\mathbf{l}^\infty(\boldsymbol{\delta})$, and $\psi^\infty(\boldsymbol{\delta})$ are defined and it is concluded as in Theorem 5.1.1 that $\boldsymbol{\delta} = \mathbf{0}$ minimizes $\psi^\infty(\boldsymbol{\delta})$. It follows that the first order conditions (14.5.13) hold at $\mathbf{x}^\infty$. In case (ii) it is also possible to conclude that second order conditions hold. $\square$

Note that the existence of a bounded region B which the theorem requires is implied if any level set $\{\mathbf{x}:\phi(\mathbf{x}) \leqslant \phi^{(k)}\}$ is bounded. Also the theorem assumes that the sequence $\{\mathbf{x}^{(k)}\}$ is infinite; if not then $\Delta\psi^{(k)} = 0$ for some k, the iteration terminates, and first order conditions are satisfied.

One point to emphasize about the theorem is that there are no hidden assumptions that certain vectors $\mathbf{a}_i^\infty$ are linearly independent or that the multipliers $\boldsymbol{\lambda}^{(k)}$ are bounded. Methods for NSO or nonlinear programming can often be proved to be convergent under such assumptions, yet can fail in practice. Thus it is important that this theorem avoids such assumptions. Another point is that $\mathbf{W}^{(k)}$ does not need to be defined as in (14.4.10) but can be any bounded matrix. Thus the theorem indicates that a corresponding quasi-Newton method, using for example $\mathbf{B}^{(k)}$ as defined in (12.3.18) in place of $\mathbf{W}^{(k)}$, can only fail if $\mathbf{B}^{(k)}$ becomes unbounded (a weaker condition on $\mathbf{B}^{(k)}$ for convergence is given by Yuan (1985a)). A weakness in the theorem requiring further thought is that it is assumed that global minima are calculated in the subproblems. This is guaranteed by convexity when $\mathbf{B}^{(k)}$ is positive definite. However, when exact second derivatives are used, $\mathbf{W}^{(k)}$ may be indefinite and it is then unrealistic to quarantee that global minima are calculated. Finally the theorem is seen to subsume the result of Madsen (1975) for the first order method like (14.4.8) in which essentially $\mathbf{W}^{(k)} = \mathbf{0}$.

It is important to consider the effect of the trust region modification on the SNQP algorithm. Clearly if the trust region bound is inactive for sufficiently large k then the satisfactory order of convergence properties of the unmodified algorithm carry through. However, Yuan (1984) gives examples which show that the trust region bound is not inactive for sufficiently large k, and the order is first order. This clearly points to the need to modify the algorithm to retain both global and second order convergence properties. As in Section 14.4 the use of an SOC step to follow the basic step suggests itself, and an appropriate subproblem for calculating the SOC step is to add a trust region bound to (14.4.34) giving

$$\begin{aligned} &\underset{\boldsymbol{\delta}}{\text{minimize}} \quad q^{(k)}(\boldsymbol{\delta}) + h([\hat{c}_i - \mathbf{a}_i^{(k)\mathrm{T}}\hat{\boldsymbol{\delta}} + \mathbf{a}_i^{(k)\mathrm{T}}\boldsymbol{\delta}]) \\ &\text{subject to} \quad \|\boldsymbol{\delta}\| \leqslant h^{(k)}. \end{aligned} \tag{14.5.19}$$

Of course if the trust region bound is inactive for sufficiently large k this algorithm retains the satisfactory order of convergence properties of the SOC algorithm. Yuan (1985b) extends this result by showing that if $\boldsymbol{\lambda}^{(k)} \to \boldsymbol{\lambda}^*$ then the trust region bound does become inactive for sufficiently large k.

In comparison with smooth unconstrained optimization (compare Theorem 5.1.2) these results are nonetheless not all that one might hope for, and it is desirable to say something about second order conditions at a limit point. To do this requires more attention to be paid to the choice of multiplier estimates $\boldsymbol{\lambda}^{(k)}$. The current rule in algorithm (14.5.6) is that if the kth iteration is successful in reducing $\phi(\mathbf{x})$, then $\boldsymbol{\lambda}^{(k+1)}$ is chosen as the multiplier vector from the kth subproblem, else $\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)}$ is set. This rule is advantageous in that it requires no extra computation, and the multipliers are bounded a priori from Lemma 14.2.1, which is important for global convergence. (It has been proposed that $\lambda^{(k)}$ should be selected as the least squares solution of the system $\mathbf{g}^{(k)} + \mathbf{A}^{(k)}\lambda = \mathbf{0}$ from (14.2.18) but this is disadvantageous on both counts. It is possible to ensure boundedness by solving a constrained least squares problem, but at the expense of a further increase in complexity.) However, it is shown by Fletcher (1982b) that the current rule may cause algorithm (14.5.6) (with or without SOC) to stick a point which, although it satisfies first order conditions (Theorem 14.2.1), fails to satisfy second order necessary conditions. Thus a second order descent direction exists at the point under consideration. This difficulty can be circumvented by the equally simple rule that $\lambda^{(k+1)}$ is always set to the multipliers of the kth subproblem, irrespective of the change in $\phi(\mathbf{x})$. This rule would seem to be undesirable when the step $\delta^{(k)}$ arising from the subproblem is large but unsuccessful. One way to compromise between these two possibilities is to devise a rule which has the property that if

$$h^{(k+1)} < \min_{j<k} h^{(j)} \tag{14.5.20}$$

(that is if $h^{(k+1)}$ is less than any previous value) then $\lambda^{(k+1)}$ is always set to the multipliers from the kth subproblem. The best choice of rule is currently under consideration.

14.6 CONSTRAINED NON-SMOOTH OPTIMIZATION

In Section 14.1 it is argued that the majority of NSO problems fall into the category of *composite NSO* problems. However, cases do arise in which it is required to minimize composite functions subject to constraints on the variables (for example, problem (14.5.2)) and such problems are the subject of this section. A suitable general format (*constrained composite NSO*) is

$$\begin{aligned} &\underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad \phi(\mathbf{x}) \triangleq f(\mathbf{x}) + h(\mathbf{c}(\mathbf{x})) \\ &\text{subject to} \quad t(\mathbf{r}(\mathbf{x})) \leqslant 0 \end{aligned} \tag{14.6.1}$$

where the objective function is the composite function (14.1.1). There is a single inequality constraint involving a composite function with $\mathbf{r}(\mathbf{x})$ ($\mathbb{R}^n \to \mathbb{R}^p$) being smooth ($\mathbb{C}^1$) and $t(\mathbf{r})$ ($\mathbb{R}^p \to \mathbb{R}$) being convex but non-smooth ($\mathbb{C}^0$). This formulation covers constraints such as $\|\mathbf{x}\|_\infty \leqslant \rho$ which appear in restricted step methods, or a general system of inequality constraints $\mathbf{r}(\mathbf{x}) \leqslant \mathbf{0}$ through the composition $t(\mathbf{r}) = \max_i r_i$. Equality constraints $r_i(\mathbf{x}) = 0$ can be included by writing them as $r_i(\mathbf{x}) \leqslant 0$ and $-r_i(\mathbf{x}) \leqslant 0$. It also includes constraints like $\|\mathbf{r}(\mathbf{x})\|_1 \leqslant \rho$ which are not conveniently handled by (7.1.1) say. Thus (14.6.1) represents a very wide range of practical optimization problems.

In practice, those applications in which $\mathbf{r}(\mathbf{x})$ is nonlinear can be handled satisfactorily by using an exact penalty function and including penalty terms derived from $\mathbf{r}(\mathbf{x})$ in the objective function. Thus most practical interest in (14.6.1) arises when $\mathbf{r}(\mathbf{x})$ is affine and $t(\mathbf{r})$ is polyhedral. A suitable type of method for solving such problems is to modify the SNQP method described in Sections 14.4 and 14.5. Thus a sequence of NQP subproblems is solved, only differing from (14.5.2) in the addition of the polyhedral constraint $t(\mathbf{r}(\mathbf{x})) \leqslant 0$. These modified subproblems can be solved by QP-like techniques modelled on those mentioned at the end of Section 12.4. The global and local convergence properties of the modified SNQP method are therefore likely to be analogous to those for the SNQP method itself (Section 14.5) although to my knowledge these have not been investigated.

The main aim of this section is to give a comprehensive presentation of local optimality conditions for (14.6.1) which reduce to those for constrained smooth optimization (Section 9.2) when $h(\mathbf{c}(\mathbf{x}))$ is not present ($m = 0$), or to those for unconstrained non-smooth optimization (Section 14.2) when $t(\mathbf{r}(\mathbf{x}))$ is not present ($p = 0$). This can be done without losing any of the essential details of the simpler cases. Thus there appear in what follows analogues of feasible direction sets, constraint qualification, an independence assumption and a separating hyperplane lemma, leading to a statement of first order necessary conditions. Both necessary and sufficient second order conditions are also set out together with a study of the regular case in which these conditions are almost necessary *and* sufficient. Finally some examples of the various conditions in action are given. The approach is based on that of Fletcher and Watson (1980) but with some extra features to handle the more general form of the constraint in (14.6.1).

Firstly the concept of feasible directions for the constraint $t(\mathbf{r}(\mathbf{x})) \leqslant 0$ is considered, using the concept of a directional sequence defined in (9.2.1) Thus at a feasible point $\mathbf{x}'$ the set of feasible directions

$$\mathscr{F}' = \{\mathbf{s} : \mathbf{s} \neq \mathbf{0}, \exists \{\mathbf{x}^{(k)}\}, t(\mathbf{r}(\mathbf{x}^{(k)})) \leqslant 0, \mathbf{x}^{(k)} \to \mathbf{x}', \mathbf{s}^{(k)} \to \mathbf{s}, \delta^{(k)} \downarrow 0\} \tag{14.6.2}$$

can be defined. This is related to the set

$$F' = \{\mathbf{s} : \mathbf{s} \neq \mathbf{0}, t' = 0 \Rightarrow \max_{\mathbf{u} \in \partial t'} \mathbf{s}^{\mathrm{T}} \mathbf{R}' \mathbf{u} \leqslant 0\} \tag{14.6.3}$$

where $\mathbf{R} = \nabla \mathbf{r}^{\mathrm{T}}$, which can be regarded as the set of feasible directions for the linearized constraint at $\mathbf{x}'$. It is very convenient if the sets $\mathscr{F}'$ and F' are the same so it is important to consider the extent to which this is true. The mechanism of linearization is seen in the following lemma which establishes an inclusion.

Lemma 14.6.1

$$F' \supseteq \mathscr{F}'$$

Proof

Let $\mathbf{s} \in \mathscr{F}'$ so that there exists a directional sequence $\mathbf{x}^{(k)} \to \mathbf{x}'$ such that $\mathbf{s}^{(k)} \to \mathbf{s}$. A Taylor series about $\mathbf{x}'$ gives

$$\mathbf{r}^{(k)} = \mathbf{r}' + \delta^{(k)} \mathbf{R}'^{\mathrm{T}} \mathbf{s}^{(k)} + o(\delta^{(k)})$$

so that $\mathbf{r}^{(k)} \to \mathbf{r}'$ is a directional sequence in $\mathbb{R}^p$ with $(\mathbf{r}^{(k)} - \mathbf{r}')/\delta^{(k)} \to \mathbf{R}'^{\mathrm{T}}\mathbf{s}$. Thus by Lemma 14.2.5 applied to $t(\mathbf{r})$

$$\max_{\mathbf{u}\in\partial t'} \mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u} = \lim \frac{t^{(k)} - t'}{\delta^{(k)}} \leqslant 0$$

if $t' = 0$ by feasibility (t' denotes $t(\mathbf{r}(\mathbf{x}'))$ etc.). Thus $\mathbf{s} \in F'$. $\square$

Unfortunately a result going the other way is not true, e.g. $t(\mathbf{r}(\mathbf{x})) = \max(x_2 - x_1^3, \ -x_2)$ as illustrated in Figure 9.2.1. However, this is an unlikely situation and usually the *constraint qualification* $\mathscr{F}' = F'$ can be assumed to hold. Indeed this result can be guaranteed under mild conditions as the following lemma shows. This uses the notation analogous to that in (14.2.30) that if $\partial t'$ has dimension q' and $\mathbf{u}' \in \partial t'$ is arbitrary, then there exists a matrix $\mathbf{F}' \in \mathbb{R}^{p \times q'}$ with columns $\mathbf{f}'_i$ $i = 1, 2, \ldots, q'$ which provide a basis for $\partial t' - \mathbf{u}'$, in that

$$\partial t' = \{\mathbf{u}\colon \mathbf{u} = \mathbf{u}' + \mathbf{F}'\mathbf{v},\ \mathbf{v} \in V' \subset \mathbb{R}^{q'}\}. \tag{14.6.4}$$

Lemma 14.6.2

Sufficient conditions for $\mathscr{F}' = F'$ at a feasible point $\mathbf{x}'$ are

(i) $t' < 0$

or if $t' = 0$ and the function $t(\mathbf{r})$ is locally linear about $\mathbf{r}'$, then either

(ii) rank $(\mathbf{R}'[\mathbf{u}' \vdots \mathbf{F}']) = q' + 1$, or

(iii) *the function* $\mathbf{r}(\mathbf{x})$ *is affine.*

Proof

If $t' < 0$ then $\mathscr{F}' = \mathbb{R}^n \backslash \mathbf{0}$ and the result is trivial, so subsequently assume that $t' = 0$. Because $\mathscr{F}' \subseteq F'$ it is only necessary to establish the reverse inclusion, so let $\mathbf{s} \in F'$. If $\max_{\mathbf{u}\in\partial t'} \mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u} < 0$ then take $\mathbf{x}^{(k)} = \mathbf{x}' + \theta^{(k)}\mathbf{s}$ for any sequence $\theta^{(k)} \downarrow 0$ and it follows that $t^{(k)} \leqslant 0$ for sufficiently large k and hence $\mathbf{s} \in \mathscr{F}'$ (if not $\exists$ a subsequence for which $t^{(k)} > 0$: using a Taylor expansion and Lemma 14.2.5 it follows that $\max_{\mathbf{u}\in\partial t'} \mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u} \geqslant 0$ which is a contradiction). If $\max_{\mathbf{u}\in\partial t'} \mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u} = 0$, let $\mathbf{u}' \in \partial t'$ be any vector for which $\mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u}' = 0$. Without loss of generality $\mathbf{u}'$ can be regarded as the arbitrary vector in (14.6.4) above. Also define

$$\partial t'_s = \{\mathbf{u} \in \partial t' \colon \mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u}' = 0\},$$

let the dimension of $\partial t'_s$ be q'_s ($q'_s < q'$), and without loss of generality let $\mathbf{f}'_i$ $i = 1, \ldots, q'_s$ be a basis for $\partial t'_s - \mathbf{u}'$. Hence

$$\mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{f}'_i \begin{cases} = 0 & i = 1, 2, \ldots, q'_s \\ < 0 & i = q'_s + 1, \ldots, q'. \end{cases}$$

If $q'_s + 1 = n$ it follows that $\mathbf{s}^{\mathrm{T}}\mathbf{R}'[\mathbf{u}' \vdots \mathbf{F}'] = \mathbf{0}^{\mathrm{T}}$ and hence that $\mathbf{s} = \mathbf{0}$ by virtue of the rank assumption, which contradicts $\mathbf{s} \in F'$. If $q'_s + 1 < n$ it is possible following the construction of Lemma 14.2.7 (with $\mathbf{D}^*$ replaced by $[\mathbf{u}' \vdots \mathbf{F}']$) to determine a smooth arc $\mathbf{x}(\theta)$ $\theta \in [0, \bar{\theta})$ for which $\mathbf{x}(0) = \mathbf{x}'$, $\dot{\mathbf{x}}(0) = \mathbf{s}$, and

$$(\mathbf{r}(\mathbf{x}(\theta)) - \mathbf{r}')^{\mathrm{T}}\mathbf{u}' = \theta\mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u}' = 0$$

$$(\mathbf{r}(\mathbf{x}(\theta)) - \mathbf{r}')^{\mathrm{T}}\mathbf{f}'_i = \theta\mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{f}'_i \begin{cases} = 0 & i = 1, 2, \ldots, q'_s, \\ < 0 & i = q'_s + 1, \ldots, q'. \end{cases}$$

It follows that

$$\max_{\mathbf{u} \in \partial t'} (\mathbf{r}(\mathbf{x}(\theta)) - \mathbf{r}')^{\mathrm{T}}\mathbf{u} = 0. \tag{14.6.5}$$

Then using the definition (14.2.29) of a locally linear function it follows that there exists a neighbourhood of $\mathbf{r}'$ such that $t(\mathbf{r}(\mathbf{x}(\theta))) = 0$, and taking any sequence $\theta^{(k)} \downarrow 0$ gives a directional sequence with $\mathbf{s} \in \mathscr{F}'$.

Finally in case (iii), if $\mathbf{r}(\mathbf{x})$ is affine, it follows that the ray $\mathbf{x}(\theta) = \mathbf{x}' + \theta\mathbf{s}$ has $\mathbf{r}(\mathbf{x}(\theta)) = \mathbf{r}' + \theta\mathbf{R}'^{\mathrm{T}}\mathbf{s}$. Thus it is possible to deduce (14.6.5) in the above directly from $\max_{\mathbf{u} \in \partial t'} \mathbf{s}^{\mathrm{T}}\mathbf{R}'\mathbf{u} = 0$, leading again to the conclusion that $\mathbf{s} \in \mathscr{F}'$. $\square$

Note that this result is somewhat stronger than that of Fletcher and Watson (1980) in which the assumption rank $\mathbf{R}' = p$ restricts the range of $\mathbf{r}(\mathbf{x})$ to have dimension $p \leqslant n$. Also Lemma 14.6.2 allows a constraint like $\|\mathbf{r}(\mathbf{x})\|_1 \leqslant \rho$ which cannot usually be handled effectively by using Lemma 9.2.2 and a polyhedral expansion of $\|\mathbf{r}\|_1$, because of the degenerate nature of the resulting constraint set.

In moving on to discuss necessary conditions at a local solution of (14.6.1) it is convenient to define the set of *descent directions* at $\mathbf{x}'$

$$\mathscr{D}(\mathbf{x}') = \mathscr{D}' = \left\{ \mathbf{s} \colon \max_{\boldsymbol{\lambda} \in \partial h'} \mathbf{s}^{\mathrm{T}}(\mathbf{g}' + \mathbf{A}'\boldsymbol{\lambda}) < 0 \right\}.$$

Then the most basic necessary condition is the following.

Lemma 14.6.3

If $\mathbf{x}^*$ *is a local minimizer then* $\mathscr{F}^* \cap \mathscr{D}^* = \varnothing$ (no feasible descent directions).

Proof

Let $\mathbf{s} \in \mathscr{F}^*$ so that there exists a feasible directional sequence $\mathbf{x}^{(k)} \to \mathbf{x}^*$ with $\mathbf{s}^{(k)} \to \mathbf{s}$. By Taylor expansion about $\mathbf{x}^*$

$$f^{(k)} = f^* + \delta^{(k)}\mathbf{g}^{*\mathrm{T}}\mathbf{s}^{(k)} + o(\delta^{(k)})$$
$$\mathbf{c}^{(k)} = \mathbf{c}^* + \delta^{(k)}\mathbf{A}^{*\mathrm{T}}\mathbf{s}^{(k)} + o(\delta^{(k)}).$$

By local optimality, $\phi^{(k)} \geqslant \phi^*$ for sufficiently large k and hence

$$0 \leqslant \frac{\phi^{(k)} - \phi^*}{\delta^{(k)}} = \frac{f^{(k)} - f^*}{\delta^{(k)}} + \frac{h(\mathbf{c}^{(k)}) - h(\mathbf{c}^*)}{\delta^{(k)}}$$

Taking the limit as $k \to \infty$, and using Lemma 14.2.5 and the fact that $\mathbf{c}^{(k)} \to \mathbf{c}^*$

is a directional sequence with direction $\mathbf{A}^{*\mathrm{T}}\mathbf{s}$, it follows that

$$0 \leqslant \max_{\lambda \in \partial h'} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}).$$

This contradicts $\mathbf{s} \in \mathscr{D}^*$ and so proves the Lemma. □

For the reasons set out in Section 9.2 it is not possible to proceed further without making a regularity assumption

$$\mathscr{F}^* \cap \mathscr{D}^* = F^* \cap \mathscr{D}^* \tag{14.6.6}$$

However this is a very weak assumption and is certainly implied by the mild sufficient conditions for constraint qualification ($\mathscr{F}^* = F^*$) of Lemma 14.6.2. With this assumption the necessary condition of Lemma 14.6.3 becomes $F^* \cap \mathscr{D}^* = \varnothing$. It is now possible to relate this condition to the existence of multipliers in Theorem 14.6.1 below. A preliminary lemma is required which extends the concept of a separating hyperplane.

Lemma 14.6.4

If, in $\mathbb{R}^n$, C is a closed convex cone, B is a non-empty compact convex set and $B \cap C = \varnothing$, then there exists hyperplane $\mathbf{s}^{\mathrm{T}}\mathbf{x} = 0$ which separates B and C.

Proof

By construction. For each point $\mathbf{b} \in B$ the closest point in C can be determined as in Lemma 9.2.5. Because B is a compact there exist points $\hat{\mathbf{b}} \in B$ and $\hat{\mathbf{g}} \in C$ which minimize $\|\mathbf{b} - \mathbf{g}\|_2$ over all $\mathbf{b} \in B$ and $\mathbf{g} \in C$. Hence for any $\mathbf{b} \in B$

$$\|\hat{\mathbf{b}} - \hat{\mathbf{g}} + \theta(\mathbf{b} - \hat{\mathbf{b}})\|_2^2 \geqslant \|\hat{\mathbf{b}} - \hat{\mathbf{g}}\|_2^2 \qquad \theta \in [0, 1]$$

and taking the limit $\theta \downarrow 0$ it follows that

$$(\mathbf{b} - \hat{\mathbf{b}})^{\mathrm{T}}(\hat{\mathbf{g}} - \hat{\mathbf{b}}) \leqslant 0.$$

But by Lemma 9.2.5, if $\mathbf{s} = \hat{\mathbf{g}} - \hat{\mathbf{b}}$ then $\mathbf{s}^{\mathrm{T}}\mathbf{x} \geqslant 0$ for all $\mathbf{x} \in C$ and $\mathbf{s}^{\mathrm{T}}\hat{\mathbf{b}} < 0$. It follows that $\mathbf{s}^{\mathrm{T}}\mathbf{b} < 0$ for all $\mathbf{b} \in B$ thus establishing the lemma. □

In stating the first main result an appropriate Lagrangian function

$$\mathscr{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{u}, \pi) = f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{c}(\mathbf{x}) + \pi\mathbf{u}^{\mathrm{T}}\mathbf{r}(\mathbf{x}) \tag{14.6.7}$$

is introduced.

Theorem 14.6.1 (First order necessary conditions)

If $\mathbf{x}^*$ locally solves (14.6.1) *and the regularity condition* (14.6.6) *holds, then there exist multipliers $\boldsymbol{\lambda}^* \in \partial h^*$, $\mathbf{u}^* \in \partial t^*$ and $\pi^* \geqslant 0$ such that*

$$t^* \leqslant 0 \qquad \text{(feasibility)}$$

$$\pi^* t^* = 0 \qquad \text{(complementarity)}$$

and

$$\begin{aligned}\mathbf{0} &= \nabla\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*)\\ &= \mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* + \pi^*\mathbf{R}^*\mathbf{u}^*\end{aligned} \tag{14.6.8}$$

Proof

The result essentially utilizes a general form of Farkas' lemma that $F^* \cap \mathscr{D}^* = \emptyset$ iff the conditions of the theorem hold. Clearly if the latter hold and $\mathbf{s} \in F^*$ then

$$t^* < 0 \Rightarrow \pi^* = 0 \Rightarrow \mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}$$

and

$$t^* = 0 \Rightarrow \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u}^* \leqslant 0 \Rightarrow \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^*) \geqslant 0$$

from (14.6.3) and (14.6.8). In both cases $\mathbf{s} \in \mathscr{D}^*$ is contradicted so that $F^* \cap \mathscr{D}^* = \emptyset$. Conversely if the conditions of the theorem do not hold then a direction $\mathbf{s} \in F^* \cap \mathscr{D}^*$ can be constructed as follows. The conditions of the theorem are equivalent to the statement that the closed convex cone

$$\begin{aligned}C = \{\mathbf{y}\colon & t^* < 0 \Rightarrow \mathbf{y} = \mathbf{0}\\ & t^* = 0 \Rightarrow \mathbf{y} = -\pi\mathbf{R}^*\mathbf{u} \quad \forall \pi \geqslant 0, \quad \mathbf{u} \in \partial t^*\},\end{aligned}$$

which is convex by convexity of ∂t^*, and the compact convex set

$$B = \{\mathbf{b}\colon \mathbf{b} = \mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda} \quad \forall \boldsymbol{\lambda} \in \partial h^*\}$$

have a common point (from (14.6.8)). Therefore if there is no common point it follows from Lemma 14.6.4 that there exists a direction $\mathbf{s}$ such that $\max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) < 0$, that is $\mathbf{s} \in \mathscr{D}^*$, and $t^* = 0 \Rightarrow \max_{\mathbf{u} \in \partial t^*} \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} \leqslant 0$, that is $\mathbf{s} \in F^*$. The theorem is then a consequence of Lemma 14.6.3 and assumption (14.6.6). □

It is also possible to analyse the effect of second order changes within the framework of (14.6.1) if the additional assumption is made that $\mathbf{c}(\mathbf{x})$ and $\mathbf{r}(\mathbf{x})$ are $\mathbb{C}^2$ functions (but $h(\mathbf{c})$ and $t(\mathbf{r})$ remain $\mathbb{C}^0$ and convex). As in Sections 9.2 and 14.2 the first step is to define certain sets of feasible directions related to directions of zero slope, so that second order effects become important. In general let $\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*$ satisfy the first order conditions of Theorem 14.6.1 and consider the set

$$\begin{aligned}X = \{\mathbf{x}\colon & \mathbf{h}(\mathbf{c}(\mathbf{x})) = h^* + (\mathbf{c}(\mathbf{x}) - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda}^*,\\ & t(\mathbf{r}(\mathbf{x})) \leqslant 0,\\ & \pi^*(\mathbf{r}(\mathbf{x}) - \mathbf{r}^*)^{\mathrm{T}}\mathbf{u}^* = 0\}.\end{aligned} \tag{14.6.9}$$

Define $\mathscr{G}^*$ as the set of normalized feasible directions with respect to X at $\mathbf{x}^*$, that is

$$\mathscr{G}^* = \{\mathbf{s}\colon \|\mathbf{s}\|_2 = 1, \quad \exists\,\{\mathbf{x}^{(k)}\}, \mathbf{x}^{(k)} \in X, \mathbf{x}^{(k)} \to \mathbf{x}^*, \mathbf{s}^{(k)} \to \mathbf{s}, \delta^{(k)} \downarrow 0\}. \tag{14.6.10}$$

This set is closely related to the set

$$G^* = \left\{ \mathbf{s}\colon \|\mathbf{s}\|_2 = 1, \quad \mathbf{s} \in F^*, \quad \max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^* \boldsymbol{\lambda}) = 0, \quad \pi^* \max_{\mathbf{u} \in \partial t^*} \mathbf{s}^{\mathrm{T}} \mathbf{R}^* \mathbf{u} = 0 \right\} \tag{14.6.11}$$

which can be interpreted as the set of linearized feasible directions that have zero slope with regard to both $\phi(\mathbf{x})$ and (if $\pi^* > 0$) $t(\mathbf{r}(\mathbf{x}))$. The extent to which $\mathscr{G}^*$ and G^* correspond is important and it is shown by linearizing the functions which define $\mathscr{G}^*$ that $\mathscr{G}^*$ is a subset of G^*.

Lemma 14.6.5

$$\mathscr{G}^* \subseteq G^*$$

Proof

Let $\mathbf{s} \in \mathscr{G}^*$ which implies $\mathbf{s} \in \mathscr{F}^*$ and hence $\mathbf{s} \in F^*$. As in Lemma 14.2.6 it follows that $\max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^* \boldsymbol{\lambda}) = 0$. By a similar argument, if $\pi^* > 0$ (and $t^* = 0$) it follows that

$$0 = \lim \frac{(\mathbf{r}^{(k)} - \mathbf{r}^*)^{\mathrm{T}} \mathbf{u}^*}{\delta^{(k)}} = \mathbf{s}^{\mathrm{T}} \mathbf{R}^* \mathbf{u}^*$$

and because $\mathbf{s} \in F^*$ it follows from (14.6.3) that $\pi^* \max_{\mathbf{u} \in \partial t^*} \mathbf{s}^{\mathrm{T}} \mathbf{R}^* \mathbf{u} = 0$. Therefore $\mathbf{s} \in G^*$. $\square$

As in Section 14.2 it is not generally possible to derive the reverse inclusion, but there are important special cases where this can be done, associated with locally linear functions. Further discussion of this is given later in this section: at present the regularity assumption

$$\mathscr{G}^* = G^* \tag{14.6.12}$$

is made (which depends on $\boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*$ if these elements are not unique). It is now possible to state and prove the second order conditions: as usual the regularity assumption is needed only in the necessary conditions.

Theorem 14.6.2 (Second order necessary conditions)

If $\mathbf{x}^*$ *locally solves* (14.6.1) *and* (14.6.6) *holds then Theorem* 14.6.1 *is valid. For each triple* $\boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*$ *which thus exists, if* (14.6.12) *holds, then*

$$\mathbf{s}^{\mathrm{T}} \nabla^2 \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) \mathbf{s} \geq 0 \quad \forall \mathbf{s} \in G^*. \tag{14.6.13}$$

Proof

Let $\mathbf{s} \in G^*$. Then $\mathbf{s} \in \mathscr{G}^*$ and $\exists$ a directional sequence feasible in (14.6.9). A Taylor expansion of $\mathscr{L}(\mathbf{x}, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*)$ about $\mathbf{x}^*$ yields

$$\mathscr{L}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) = \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) + \mathbf{e}^{(k)\mathrm{T}} \nabla \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) + \tfrac{1}{2} \mathbf{e}^{(k)\mathrm{T}} \nabla^2 \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) \mathbf{e}^{(k)} + o(\|\mathbf{e}^{(k)}\|^2) \tag{14.6.14}$$

where $\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^*$. The definition of $\mathscr{L}$ yields

$$\begin{aligned}
&\mathscr{L}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) - \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) \\
&\quad = f^{(k)} - f^* + (\mathbf{c}^{(k)} - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda}^* + \pi^*(\mathbf{r}^{(k)} - r^*)^{\mathrm{T}}\mathbf{u}^* \\
&\quad = f^{(k)} - f^* + h(\mathbf{c}^{(k)}) - h^* = \phi^{(k)} - \phi^*
\end{aligned} \tag{14.6.15}$$

by feasibility in (14.6.9). Substituting $\mathbf{e}^{(k)} = \delta^{(k)}\mathbf{s}^{(k)}$ and the local minimality of ϕ^* then gives

$$0 \leqslant \phi^{(k)} - \phi^* = \tfrac{1}{2}\delta^{(k)^2}\mathbf{s}^{(k)\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*)\mathbf{s}^{(k)} + o(\delta^{(k)^2}).$$

Finally dividing by $\frac{1}{2}\delta^{(k)^2}$ and taking the limit establishes (14.6.13). □

Theorem 14.6.3 (Second order sufficient conditions)

If at $\mathbf{x}^*, t^* \leqslant 0$ *and there exist* $\boldsymbol{\lambda}^* \in \partial h^*$, $\mathbf{u}^* \in \partial t^*$ *and* $\pi^* \geqslant 0$ *such that* $\pi^* t^* = 0$ *and* (14.6.8) *holds (first order conditions), and if*

$$\mathbf{s}^{\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*)\mathbf{s} > 0 \qquad \forall\, \mathbf{s} \in G^* \tag{14.6.16}$$

then $\mathbf{x}^*$ *is a strict local solution of* (14.6.1).

Proof

Assume the contrary that $\exists$ a feasible sequence and hence a feasible direction sequence $\mathbf{x}^{(k)} \to \mathbf{x}^*$, $\mathbf{s}^{(k)} \to \mathbf{s}$, $\|\mathbf{s}\|_2 = 1$ such that $\phi^{(k)} \leqslant \phi^*$. Now $\mathbf{s} \in \mathscr{F}^* \Rightarrow \mathbf{s} \in F^*$, so using (14.6.8),

$$t^* < 0 \Rightarrow \pi^* = 0 \Rightarrow \mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^* = \mathbf{0}$$

and from $\mathbf{s} \in F^*$

$$t^* = 0 \Rightarrow \max_{\mathbf{u} \in \partial t^*} \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} \leqslant 0$$

so in both cases it follows that

$$0 \leqslant \max_{\boldsymbol{\lambda} \in \partial h^*} \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) = \mu \quad \text{say.}$$

If $\mu > 0$ then from Lemma 14.2.5, $\lim(\phi^{(k)} - \phi^*)/\delta^{(k)} = \mu$ which contradicts $\phi^{(k)} \leqslant \phi^*$. Thus $\mu = 0$. Let $\mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^*) < 0$; then the first order conditions imply $\pi^* > 0$, $t^* = 0$ and $\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u}^* > 0$ which contradicts $\mathbf{s} \in F^*$. Thus $\mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^*) = 0$ and hence $\pi^*\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u}^* = 0$. It follows from $\mathbf{s} \in F^*$ that $\pi^* \max_{\mathbf{u} \in \partial t^*} \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u}^* = 0$ and hence $\mathbf{s} \in G^*$. Now from (14.6.15)

$$\begin{aligned}
&\mathscr{L}(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) - \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*) \\
&\quad = \phi^{(k)} - \phi^* - (h^{(k)} - h^* - (\mathbf{c}^{(k)} - \mathbf{c}^*)^{\mathrm{T}}\boldsymbol{\lambda}^*) \\
&\qquad + \pi^*(t^{(k)} - t^* - (t^{(k)} - t^* - (\mathbf{r}^{(k)} - \mathbf{r}^*)^{\mathrm{T}}\mathbf{u}^*)) \\
&\quad \leqslant \phi^{(k)} - \phi^* + \pi^*(t^{(k)} - t^*) \leqslant \phi^{(k)} - \phi^*
\end{aligned}$$

using the subgradient inequality and then feasibility. Hence from (14.6.14) and (14.6.8)

$$0 \geqslant \phi^{(k)} - \phi^* \geqslant \tfrac{1}{2}\delta^{(k)^2}\mathbf{s}^{(k)\mathrm{T}}\nabla^2\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*, \pi^*)\mathbf{s}^{(k)} + o(\delta^{(k)^2}).$$

Dividing by $\frac{1}{2}\delta^{(k)^2}$ and taking the limit contradicts (14.6.16) and establishes the theorem. □

Corollary

If the directional derivative $\max_{\boldsymbol{\lambda}\in\partial h^*}\mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda})$ *is positive for all feasible directions in* F^*, *or equivalently if* G^* *is empty, then the first order conditions are sufficient to imply that* $\mathbf{x}^*$ *is a strict local solution of* (14.6.1).

Proof

Immediate from the statement of Theorem 14.6.2. □

The remarks after Theorem 14.2.3 about when there is near equivalence in the second order conditions apply equally here. The conditions are 'almost' necessary and sufficient only if it can be shown that $\mathscr{G}^* = G^*$ and this result is only true for particular types of composite function having the locally linear property. Effectively either $h(\mathbf{c})$ and $t(\mathbf{r})$ are restricted to being polyhedral, or $h(\mathbf{c})$ can be the penalty term in certain exact penalty function applications (Lemma 14.3.1). A mild independence assumption then enables $\mathscr{G}^* = G^*$ to be established using the constructions of (14.2.30) and (14.6.4) as follows.

Lemma 14.6.6 (Sufficient conditions for regularity)

If $\mathbf{x}^*$ *satisfies the first order conditions of Theorem* 14.6.1, *if* $h(\mathbf{c})$ *and* $t(\mathbf{r})$ *are* locally linear at $\mathbf{c}^*$ *and* $\mathbf{r}^*$ *and if*

$$\operatorname{rank}([\mathbf{A}^*\mathbf{D}^* : \mathbf{R}^*\mathbf{u}^* : \mathbf{R}^*\mathbf{F}^*]) = l^* + q^* + 1 \tag{14.6.17}$$

then $\mathscr{G}^* = G^*$. *Alternatively the rank assumption may be replaced by an assumption that the functions* $\mathbf{c}(\mathbf{x})$ *and* $\mathbf{r}(\mathbf{x})$ *are affine.*

Proof

Because $G^* \supseteq \mathscr{G}^*$ it is sufficient to establish the reverse inclusion. Let $\mathbf{s}\in G^*$. If $t^* < 0$ or if $t^* = 0$, $\pi^* = 0$ and $\max_{\mathbf{u}\in\partial t^*}\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} < 0$ then an arc and hence a directional sequence is constructed as in Lemma 14.2.7. This sequence is easily shown to be feasible ($t^{(k)} \leqslant 0$): in the latter case the argument of Lemma 14.6.2 is used. Thus $\mathbf{s}\in\mathscr{G}^*$. If $\pi^* = 0$ and $\max_{\mathbf{u}\in\partial t^*}\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} = 0$ then without loss of generality $\mathbf{u}^*$ can be taken to attain the max, so that $\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u}^* = 0$. This enables this case and the general case $\pi^* > 0$ to be treated together, as follows. Define

$$\partial h_s^* = \{\boldsymbol{\lambda} \in \partial h^* : \mathbf{s}^{\mathrm{T}}(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}) = 0\}$$
$$\partial t_s^* = \{\mathbf{u} \in \partial t^* : \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} = 0\}$$

which depend on $\mathbf{s}$. Now $\mathbf{s} \in G^*$ and the first order conditions imply that

$$\mathbf{s}^T(\mathbf{g}^* + \mathbf{A}^*\boldsymbol{\lambda}^*) = \pi^*\mathbf{s}^T\mathbf{R}^*\mathbf{u}^* = 0$$

so that

$$\begin{aligned} &\mathbf{s}^T\mathbf{A}^*(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) = 0 \qquad && \forall \boldsymbol{\lambda} \in \partial h_s^* \\ &\mathbf{s}^T\mathbf{R}^*\mathbf{u} = 0 && \forall \mathbf{u} \in \partial t_s^* \end{aligned}$$

and from (14.6.11)

$$\mathbf{s}^T\mathbf{A}^*(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) < 0 \quad \forall \boldsymbol{\lambda} \in \partial h^* \backslash \partial h_s^*$$

$$\mathbf{s}^T\mathbf{R}^*\mathbf{u} < 0 \quad \forall \mathbf{u} \in \partial t^* \backslash \partial t_s^*$$

Let the dimensions of ∂h_s^* and ∂t_s^* be l_s^* and q_s^* respectively and let $\mathbf{u}^*$ be the arbitrary vector $\mathbf{u}'$ in (14.6.4). Assume without loss of generality that the vectors $\mathbf{d}_i^*$ $i = 1, \ldots, l_s^*$ and $\mathbf{f}_i^*$ $i = 1, \ldots, q_s^*$ form bases for $\partial h_s^* - \boldsymbol{\lambda}^*$ and $\partial t_s^* - \mathbf{u}^*$ respectively. Then

$$\mathbf{s}^T\mathbf{A}^*\mathbf{d}_i^* \begin{cases} = 0 & i = 1, 2, \ldots, l_s^* \\ < 0 & i = l_s^* + 1, \ldots, l^* \end{cases}$$

$$\mathbf{s}^T\mathbf{R}^*\mathbf{u}^* = 0$$

$$\mathbf{s}^T\mathbf{R}^*\mathbf{f}_i^* \begin{cases} = 0 & i = 1, 2, \ldots, q_s^* \\ < 0 & i = q_s^* + 1, \ldots, q^* \end{cases}$$

If $l_s^* + q_s^* + 1 = n$ then it follows from the rank assumption that $\mathbf{s} = \mathbf{0}$ which contradicts $\mathbf{s} \in G^*$. For $l_s^* + q_s^* + 1 < n$ it is possible as in Lemma 14.2.7 to construct a smooth arc $\mathbf{x}(\theta)$ $\theta \in [0, \bar{\theta})$ for which $\mathbf{x}(0) = \mathbf{x}^*$, $\dot{\mathbf{x}}(0) = \mathbf{s}$ and

$$\begin{aligned} &(\mathbf{c}(\mathbf{x}(\theta)) - \mathbf{c}^*)^T\mathbf{d}_i^* = \theta\mathbf{s}^T\mathbf{A}^*\mathbf{d}_i^* \qquad && i = 1, 2, \ldots, l^* \\ &(\mathbf{r}(\mathbf{x}(\theta)) - \mathbf{r}^*)^T\mathbf{u}^* = \theta\mathbf{s}^T\mathbf{R}^*\mathbf{u}^* && \\ &(\mathbf{r}(\mathbf{x}(\theta)) - \mathbf{r}^*)^T\mathbf{f}_i^* = \theta\mathbf{s}^T\mathbf{R}^*\mathbf{f}_i^* && i = 1, 2, \ldots, q^*. \end{aligned} \tag{14.6.18}$$

Consequently in a similar way to Lemma 14.2.7

$$\max_{\boldsymbol{\lambda} \in \partial h^*} (\mathbf{c}(\mathbf{x}(\theta)) - \mathbf{c}^*)^T(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) = 0 \tag{14.6.19}$$

$$\max_{\mathbf{u} \in \partial t^*} (\mathbf{r}(\mathbf{x}(\theta) - \mathbf{r}^*)^T\mathbf{u} = 0 \tag{14.6.20}$$

and hence from the local linearity assumption

$$h(\mathbf{c}(\mathbf{x}(\theta))) = h^* + (\mathbf{c}(\mathbf{x}(\theta)) - \mathbf{c}^*)^T\boldsymbol{\lambda}^*$$
$$t(\mathbf{r}(\mathbf{x}(\theta))) = 0.$$

Adding to these the equation

$$\pi^*(\mathbf{r}(\mathbf{x}(\theta)) - \mathbf{r}^*)^T\mathbf{u}^* = 0 \tag{14.6.21}$$

from (14.6.18), it follows that $\mathbf{x}(\theta) \in X$ in (14.6.9) and hence by taking any sequence $\theta^{(k)} \downarrow 0$, $\mathbf{s}$ is shown to be a feasible direction in $\mathscr{G}^*$.

Alternatively if the assumption is made that $\mathbf{c}(\mathbf{x})$ and $\mathbf{r}(\mathbf{x})$ are affine in place of (14.6.17), then the arc $\mathbf{x}(\theta) = \mathbf{x}^* + \theta\mathbf{s}$ has

$$\mathbf{c}(\mathbf{x}(\theta)) = \mathbf{c}^* + \mathbf{A}^{*\mathrm{T}}\mathbf{s}$$
$$\mathbf{r}(\mathbf{x}(\theta)) = \mathbf{r}^* + \mathbf{R}^{*\mathrm{T}}\mathbf{s}$$

and it is possible to deduce (14.6.19), (14.6.20) and (14.6.21) directly from the equations

$$\max_{\lambda\in\partial h^*} \mathbf{s}^{\mathrm{T}}\mathbf{A}^*(\boldsymbol{\lambda} - \boldsymbol{\lambda}^*) = 0$$
$$\max_{\mathbf{u}\in\partial t^*} \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} = 0$$
$$\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u}^* = 0$$

which arise earlier in the proof. □

This section finishes with a number of remarks and examples to clarify the way in which the conditions operate. It is intrinsic in the development of the conditions that $t(\mathbf{r})$ has the property of local linearity. An example for which this is not the case is a constraint like $\|\mathbf{r}(\mathbf{x})\|_2 \leqslant \rho$ for $\rho > 0$, in which $t(\mathbf{r}) = \|\mathbf{r}\|_2$ is locally smooth and nonlinear. Such a constraint may be brought within the scope of Lemma 14.6.6 by replacing it by $|\bar{r}(\mathbf{x})| \leqslant \rho$ where $\bar{r}(\mathbf{x}) = \|\mathbf{r}(\mathbf{x})\|_2$. Thus the problem is essentially transformed to one in which $t(\bar{r})$ is a polyhedral function on $\mathbb{R}^1$. ∂t^* contains just the single element $u^* = \operatorname{sign}(\bar{r}(\mathbf{x}^*))$ and if $\pi^* > 0$ the third condition in (14.6.9) simply reduces to the requirement that $\bar{r}(\mathbf{x}) = \bar{r}(\mathbf{x}^*)$.

An important application of the conditions is the use of the single constraint $t(\mathbf{r}(\mathbf{x})) \leqslant 0$ to represent a set of inequality constraints $\mathbf{r}(\mathbf{x}) \leqslant \mathbf{0}$ using the polyhedral convex function $t(\mathbf{r}) = \max_i r_i$. The first order conditions (14.6.8) involve the term $\pi^*\mathbf{R}^*\mathbf{u}^*$ and the elements $\pi^* u_i^*$ become the Lagrange multipliers of the individual constraints $r_i(x) \leqslant 0$. If $\mathscr{A}^*$ indexes the active constraints then the definition of ∂t^* shows that $\mathbf{u}^*$ satisfies $\mathbf{u}^* \geqslant \mathbf{0}$, $\sum_{i\in\mathscr{A}^*} u_i^* = 1$ and $u_i^* = 0\; i \notin \mathscr{A}^*$ which shows that $\pi^* u_i^*$ satisfies the usual complementarity conditions for inequality constraints, and π^* is the L_1 norm of the Lagrange multiplier vector $[\pi^* u_i^*]$. The set X in (14.6.9) involves the conditions $\max_i r_i(\mathbf{x}) \leqslant 0$ and $\pi^*(\mathbf{r}(\mathbf{x}) - \mathbf{r}^*)^{\mathrm{T}}\mathbf{u}^* = 0$ which is equivalent to $\sum_{i\in\mathscr{A}^*_+} \pi^* u_i^* r_i(\mathbf{x}) = 0$ where $\mathscr{A}^*_+$ denotes active constraints with positive multipliers. This is again equivalent to $r_i(\mathbf{x}) = 0\; i \in \mathscr{A}^*_+$ which can be recognized as the condition for defining $\mathscr{G}^*$ in Section 9.3. Likewise the condition $\max_{\mathbf{u}\in\partial t^*} \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} \leqslant 0$ in F^* in (14.6.3) is equivalent to $\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{e}_i \leqslant 0$ for $i \in \mathscr{A}^*$ and the condition $\pi^* \max_{\mathbf{u}\in\partial^*} \mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{u} = 0$ is equivalent to $\mathbf{s}^{\mathrm{T}}\mathbf{R}^*\mathbf{e}_i = 0$ for $i \in \mathscr{A}^*_+$ so that these conditions which define G^* can be identified with the equivalent conditions in (9.3.11) in Section 9.3.

Finally a numerical example of the first and second order conditions is given in which the constraint involves the L_1 norm. Consider the problem

$$\begin{aligned} &\text{minimize} && \|\mathbf{x}\|_\infty \\ &\text{subject to} && \|\mathbf{r}(\mathbf{x})\|_1 \leqslant 0.5 \end{aligned} \qquad (14.6.22)$$

where $\mathbf{r}:\mathbb{R}^3 \to \mathbb{R}^4$ is defined by

$$\begin{aligned} r_1 &= x_1^2 + x_2^2 - 1 \\ r_2 &= x_1 x_2 - 0.5 \\ r_3 &= x_1 + \tfrac{1}{2}(x_2^2 - 1) \\ r_4 &= -x_1 + x_3^2 + 0.54. \end{aligned} \tag{14.6.23}$$

This problem is an example of (14.6.1) with $f(\mathbf{x}) = 0$, $\mathbf{c}(\mathbf{x}) = \mathbf{x}$, $h(\mathbf{c}) = \|\mathbf{c}\|_\infty$ and $t(\mathbf{r}) = \|\mathbf{r}\|_1 - 0.5$. At $\mathbf{x}^* = (0.6, 0.8, 0)^T$, $\mathbf{r}^* = (0, -0.02, 0.42, -0.06)^T$ and the first order conditions are satisfied by $\mathbf{A}^* = \mathbf{I}$, $\pi^* = \frac{5}{7}$ and

$$\boldsymbol{\lambda}^* = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{R}^* = \begin{bmatrix} 1.2 & 0.8 & 1 & -1 \\ 1.6 & 0.6 & 0.8 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{u}^* = \begin{pmatrix} -1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

The set $G^* = \{\mathbf{s}: \|\mathbf{s}\|_2 = 1, s_2 = 0, s_1 \leqslant 0\}$, the dimensions of ∂h^* and ∂t^* are $l^* = 0$ and $t^* = 1$ respectively, and the basis for $\partial t^* - \mathbf{u}^*$ is $\mathbf{F}^* = [1 \quad 0 \quad 0 \quad 0]^T$. Thus the matrix $[\mathbf{R}^*\mathbf{u}^* : \mathbf{R}^*\mathbf{F}^*]$ has rank 2, and since $h(\mathbf{c})$ and $t(\mathbf{r})$ are polyhedral it follows from Lemma 14.6.6 that the regularity assumption $\mathscr{G}^* = G^*$ holds. In fact the set $\mathscr{G}^*$ can be obtained explicitly as follows. The first condition that defines X in (14.6.9) is $\|\mathbf{x}\|_\infty = x_2$ which is locally vacuous. When $\pi^* > 0$ the remaining conditions imply that $t(\mathbf{r}(\mathbf{x})) = 0$, and the last condition $(\mathbf{r} - \mathbf{r}^*)^T\mathbf{u}^* = 0 = \mathbf{r}^T\mathbf{u}^* - 0.5$ implies that $\operatorname{sign} r_i = u_i^*$ and in particular that $r_1 \leqslant 1$. After some manipulation using (14.6.23) the condition $(\mathbf{r} - \mathbf{r}^*)^T\mathbf{u}^*$ yields

$$d_1^2 + \tfrac{1}{2}d_2^2 + 1.4d_2 + d_3^2 = 0$$

where $d_i = x_i - x_i^*$ $i = 1, 2, 3$. Thus $d_2 = O(d_1^2) + O(d_3^2) \leqslant 0$ and feasible directions are seen to have $s_1 = s_3 = 0$. In addition the condition $r_1 \leqslant 0$ imposes $d_1 \leqslant 0$ (for otherwise $d_1 > 0$ and $d_2 = O(d_1^2)$ gives a contradiction) so the extra condition $s_1 \leqslant 0$ is implied and hence $\mathscr{G}^* = G^*$. Because $d_2 < 0$ if d_1 or d_2 is non-zero it can be seen that $\mathbf{x}^*$ is not a minimizing point. However, this result can also be deduced from the Lagrangian condition (14.6.13) because

$$\mathbf{s}^T\nabla^2\mathscr{L}^*\mathbf{s} = -\tfrac{5}{7}\mathbf{s}^T \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \mathbf{s} < 0$$

for both $\mathbf{s} = \mathbf{e}_1$ and $\mathbf{s} = \mathbf{e}_3$, that is for all $\mathbf{s} \in G^*$. This shows that $\mathbf{x}^*$ fails to satisfy the second order necessary conditions, and hence $\mathbf{x}^*$ is not a solution since $\mathscr{F}^* = F^*$ and $\mathscr{G}^* = G^*$.

In fact the solution to the problem is given by $x_1^* = x_2^* = (1 + \sqrt{6})/5 \doteq 0.689898$ and $x_3^* = \sqrt{(x_1^* - 0.54)} \doteq 0.387167$. The multipliers $\boldsymbol{\lambda}^* = (0.436700, 0.563299)^T$, $\mathbf{u}^* = (-1, -1, 1, 0)^T$ and $\pi^* = 0.408247$ satisfy the first order conditions. Because $\boldsymbol{\lambda}^*$ and $\mathbf{u}^*$ are in the relative interiors of ∂h^* and ∂t^*, because $l^* + q^* + 1 = 3$, and because the rank condition (14.6.17) is satisfied, it follows that $\mathscr{G}^* = G^* = \varnothing$ and the first order conditions are sufficient to show that $\mathbf{x}^*$ is a solution. Another example which is a variation on this theme is given in Question 14.15.

QUESTIONS FOR CHAPTER 14

14.1. For the nonlinear program (14.1.6) prove that $\mathscr{F}' = F'$ (defined in Section 9.2). Let $\bar{\mathbf{s}} = (\mathbf{s}^{\mathrm{T}}, s_{n+1})^{\mathrm{T}}$ be a feasible direction, and use the linearized constraint equations to show that

$$s_{n+1} \geqslant \max_{i \in \mathscr{A}'} \mathbf{s}^{\mathrm{T}}(\mathbf{g}' + \mathbf{A}'\mathbf{h}_i). \tag{a}$$

Let $\delta^{(k)} \downarrow 0$ be any sequence, and define $\mathbf{x}^{(k)} = \mathbf{x}' + \delta^{(k)}\mathbf{s}$. If equality holds in (a), and if $v^{(k)} = \phi(\mathbf{x}^{(k)})$, show that $\mathbf{x}^{(k)}, v^{(k)}$ gives a feasible directional sequence in (14.1.6) for sufficiently large k. If strict inequality holds in (a), and if $v^{(k)} = v^* + \delta^{(k)} s_{n+1}$, again show that $\mathbf{x}^{(k)}$, $v^{(k)}$ is a feasible directional sequence.

14.2. Prove that $\|\mathbf{c}^+\|$ is a convex function of $\mathbf{c}$ when the norm is monotonic ($|\mathbf{x}| \leqslant |\mathbf{y}| \Rightarrow \|\mathbf{x}\| \leqslant \|\mathbf{y}\|$). Deduce that $\mathbf{c}_\theta^+ \leqslant (1-\theta)\mathbf{c}_0^+ + \theta\mathbf{c}_1^+$ as an intermediate stage.

14.3. Establish the equivalence between each of (14.1.11) to (14.1.15) and the general expression $\partial h(\mathbf{c}) = \operatorname{conv}_{i \in \mathscr{A}} \mathbf{h}_i$. In (14.1.12) let $\lambda_i = \mu_i$, $i \leqslant m$, and μ_{m+1} acts as the slack variable for $\sum_i \lambda_i \leqslant 1$. In (14.1.13) define $\lambda_i = \mu_i - \mu_{m+i}$ or $\mu_i = \max(\lambda_i, 0)$, $\mu_{m+i} = \max(-\lambda_i, 0)$. In (14.1.14) use the fact that the cube $0 \leqslant \lambda_i \leqslant 1$ has extreme points which are all combinations of 1 and 0, and similarly for (14.1.15).

14.4. Establish the equivalence between the subdifferential expression

$$\partial\|\mathbf{c}\| = \{\boldsymbol{\lambda}\colon \|\mathbf{c} + \mathbf{h}\| \geqslant \|\mathbf{c}\| + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h} \quad \forall\, \mathbf{h}\} \tag{b}$$

and (14.3.7). Use the generalized Cauchy inequality $\mathbf{a}^{\mathrm{T}}\mathbf{b} \leqslant \|\mathbf{a}\|\,\|\mathbf{b}\|_D$ on $(\mathbf{c}+\mathbf{h})^{\mathrm{T}}\boldsymbol{\lambda}$ to show that $\boldsymbol{\lambda} \in$(14.3.7) implies $\boldsymbol{\lambda} \in$(b). If $\boldsymbol{\lambda} \in$(b), use the triangle inequality to show that $\mathbf{h}^{\mathrm{T}}\boldsymbol{\lambda} \leqslant \|\mathbf{h}\|$ $\forall\, \mathbf{h}$ and (14.3.4) to show $\|\boldsymbol{\lambda}\|_D \leqslant 1$. Hence $\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{c} \leqslant \|\mathbf{c}\|$. Then with $\mathbf{h} = -\mathbf{c}$ in (b) show that $\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{c} \geqslant \|\mathbf{c}\|$ and hence $\boldsymbol{\lambda} \in$(14.3.7).

14.5. If the norm is monotonic (see Question 14.2) establish the equivalence between

$$\partial\|\mathbf{c}^+\| = \{\boldsymbol{\lambda}: \|(\mathbf{c}+\mathbf{h})^+\| \geqslant \|\mathbf{c}^+\| + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h} \quad \forall\, \mathbf{h}\} \tag{c}$$

and (14.3.8). The proof is similar to that in Question 14.4. In the first part also use the fact that $\boldsymbol{\lambda} \geqslant \mathbf{0}$ implies $(\mathbf{c}+\mathbf{h})^{+\mathrm{T}}\boldsymbol{\lambda} \geqslant (\mathbf{c}+\mathbf{h})^{\mathrm{T}}\boldsymbol{\lambda}$. In the second part also use the monotonic norm property to establish $\|\mathbf{h}^+\| \geqslant \mathbf{h}^{\mathrm{T}}\boldsymbol{\lambda}$ $\forall\, \mathbf{h}$. Then $\mathbf{h} = -\mathbf{e}_i$ yields $\lambda_i \geqslant 0$. Use $\mathbf{h}^+ \leqslant |\mathbf{h}|$ to show $\|\boldsymbol{\lambda}\|_D \leqslant 1$. Then proceed much as in Question 14.4.

14.6. Show that the subdifferential in (14.2.2) is a closed convex set.

14.7. Justify equation (14.2.12) when $\mathbf{0}\notin\partial f'$. Show by straightforward arguments that

$$\min_{\|\mathbf{s}\|_2=1}\max_{\mathbf{g}\in\partial f'}\mathbf{s}^{\mathrm{T}}\mathbf{g}\geqslant\max_{\mathbf{g}\in\partial f'}\min_{\|\mathbf{s}\|_2=1}\mathbf{s}^{\mathrm{T}}\mathbf{g}=-\|\bar{\mathbf{g}}\|_2.$$

Then use the separating hyperplane result (Lemma 14.2.3) to show that equality is achieved when $\mathbf{s}=-\bar{\mathbf{g}}/\|\bar{\mathbf{g}}\|_2$. Show by considering $f=|x|$ that the result is not true when $\mathbf{0}\in\partial f'$. The result can be generalized to include the case $\mathbf{0}\in\partial f'$ by writing $\|\mathbf{s}\|_2\leqslant 1$ in place of $\|\mathbf{s}\|_2=1$.

14.8. Consider the Freudenstein and Roth equations

$$c_1(\mathbf{x})=x_1-x_2^3+5x_2^2-\ 2x_2-13$$
$$c_2(\mathbf{x})=x_1+x_2^3+\ x_2^2-14x_2-29$$

(see Chapter 6) and consider minimizing $\|\mathbf{c}(\mathbf{x})\|_p$ for $1\leqslant p\leqslant\infty$. A local solution in all cases is $\mathbf{x}^*=(11.4128,-0.8968)^{\mathrm{T}}$. Find the sets $\partial\|\mathbf{c}^*\|_p$ (that is ∂h^*) for $p=1,2,\infty$ and the vector $\boldsymbol{\lambda}^*$ which satisfies Theorem 14.2.1. For $p=1$ the local solution is not unique and any $x_1^*\in[6.4638, 11.4128]$ gives a solution. Find $\partial\|\mathbf{c}^*\|_1$ when $x_1^*=6.4638$.

14.9. Consider minimizing $\|\mathbf{c}\|_1$ when the equation $c_1(\mathbf{x})$ in Question 14.8 is scaled by multiplying the right-hand side by 2. Show that $\mathbf{x}^*=(6.4638, -0.8968)^{\mathrm{T}}$ satsifies the second order sufficient conditions of Theorem 14.2.3 and find $\boldsymbol{\lambda}^*$. Find a trajectory which is feasible in (14.2.19) (see Figure 14.2.3) and hence find the sets $\mathscr{G}^*$ and G^* and verify that they are non-empty and equal.

14.10. Consider minimizing $\|\mathbf{c}\|_\infty$ for the system which results on adding an equation $c_3(\mathbf{x})=\alpha x_1$ to those in Question 14.8. If $\alpha=0.4336$ show that $\mathbf{x}^*=(11.4128,-0.8968)^{\mathrm{T}}$ satisfies the second order sufficient conditions of Theorem 14.2.3 and find $\boldsymbol{\lambda}^*$. Find a trajectory which is feasible in (14.2.19) and hence find the sets $\mathscr{G}^*$ and G^* and verify that they are non-empty and equal.

14.11. For the problems defined in Questions 14.9 and 14.10 derive the dimension l^* and the basis $\mathbf{D}^*$ for the set $\partial h^*-\boldsymbol{\lambda}^*$ in (14.2.30) and show that the rank assumption (14.2.31) is satisfied in both cases.

14.12. Consider the problem defined in (14.4.2). At $\mathbf{x}^\infty$ find the multipliers $\boldsymbol{\lambda}^\infty$ for which $\mathbf{g}^\infty+\mathbf{A}^\infty\boldsymbol{\lambda}^\infty=\mathbf{0}$ and show that $\lambda_2^\infty<0$ which implies that $\phi(\mathbf{x})$ can be reduced by making the inequality $v\geqslant c_2(\mathbf{x})$ in (14.1.6) inactive (see end of Section 14.1). Find the multipliers $\boldsymbol{\lambda}^*$ at $\mathbf{x}^*$ and show that the conditions of Theorem 14.2.1 hold. Apply the method of (14.4.8) from $\mathbf{x}^{(1)}=(0,-4)^{\mathrm{T}}$ and verify that second order convergence is obtained. Why does this happen in the absence of curvature information?

14.13. In problem (12.3.2) let $\mathbf{x}^{(1)}$ lie on the unit circle arbitrarily close to $\mathbf{x}^*$, and let $\boldsymbol{\lambda}^{(1)}=\boldsymbol{\lambda}^*$. Show that there exists a range of values of v with $v<1/\lambda^*$ for which the unit step determined by solving (14.4.11) fails to reduce $\phi(\mathbf{x})$ (the Maratos effect).

14.14. Consider the unconstrained NSO problem: $\min f(\mathbf{x}) + \sum_{i=1}^{m} |r_i(\mathbf{x})|$, where $\mathbf{r}(\mathbf{x}) = \mathbf{A}^T\mathbf{x} + \mathbf{b}$ $(\mathbf{x} \in \mathbb{R}^n)$ and $f(\mathbf{x})$ is convex. By introducing variables

$$r_i^+ = \max(r_i, 0), \qquad r_i^- = \max(-r_i, 0),$$

show that $r_i^+ - r_i \geqslant 0$, $r_i^- + r_i \geqslant 0$, and $r_i^+ + r_i^- = |r_i|$. Hence show that the unconstrained problem can be restated as

$$\begin{aligned} \underset{\mathbf{x},\mathbf{r}^+,\mathbf{r}^-}{\text{minimize}} \quad & f(\mathbf{x}) + \mathbf{e}^T(\mathbf{r}^+ + \mathbf{r}^-) \\ \text{subject to} \quad & \mathbf{r}^+ - \mathbf{A}^T\mathbf{x} - \mathbf{b} \geqslant \mathbf{0} \\ & \mathbf{r}^- + \mathbf{A}^T\mathbf{x} + \mathbf{b} \geqslant \mathbf{0}, \qquad \mathbf{r}^+ \geqslant \mathbf{0}, \mathbf{r}^- \geqslant \mathbf{0}, \end{aligned}$$

where $\mathbf{e} = (1, 1, \ldots, 1)^T$. Show that this is a convex programming problem. Write down the dual of this problem, denoting the multipliers of the constraints by $\boldsymbol{\lambda}^+, \boldsymbol{\lambda}^-, \boldsymbol{\mu}^+, \boldsymbol{\mu}^-$ respectively. By eliminating $\boldsymbol{\mu}^+$ and $\boldsymbol{\mu}^-$, and writing $\boldsymbol{\lambda}^+ - \boldsymbol{\lambda}^- = \boldsymbol{\lambda}$, show that the dual can be restated more simply as

$$\begin{aligned} \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{maximize}} \quad & f(\mathbf{x}) + \boldsymbol{\lambda}^T(\mathbf{A}^T\mathbf{x} + \mathbf{b}) \\ \text{subject to} \quad & \mathbf{g}(\mathbf{x}) + \mathbf{A}\boldsymbol{\lambda} = \mathbf{0}, \qquad -\mathbf{e} \leqslant \boldsymbol{\lambda} \leqslant \mathbf{e}, \end{aligned}$$

where $\mathbf{g} = \nabla_x f$ (see also Watson, 1978).

14.15. Show that the point $\mathbf{x}^* = (0.6, 0.8, 0)$ satisfies first but not second order necessary conditions for the problem

$$\begin{aligned} \text{minimize} \quad & \|\mathbf{r}(\mathbf{x})\|_1 \\ \text{subject to} \quad & \|\mathbf{x}\|_\infty \leqslant 0.8 \end{aligned}$$

where $\mathbf{r}(\mathbf{x})$ is defined by (14.6.23). Find the solution of the problem.

14.16. Apropos of Theorem 14.4.1, construct a problem satisfying the conditions of the theorem such that the SNQP subproblem as defined at $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ does not have $\boldsymbol{\delta} = \mathbf{0}$ and $\boldsymbol{\lambda}^+ = \boldsymbol{\lambda}^*$ as its global minimizer. (Hint: choose $\mathbf{W}^*$ to be indefinite.)