HYBRID DIRECTION METHOD FOR SOLVING UNCONSTRAINED MINIMIZATION PROBLEMS

Prapasri Asawakun[1] and Phaichayon Sirisathienwatthana[2]

[1] *School of Mathematics, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, Thailand*
[2] *Department of Mathematics, Rajabhat Kumphaengphet Institute, Kumphaengphet, Thailand*

ABSTRACT: Various search directions have been applied to find a minimizer in an unconstrained minimization problem, such as steepest descent direction, Newton direction, quasi-Newton directions, conjugate gradient direction, coordinate directions, etc. In the present investigation, some of these directions are linearly combined to produce a hybrid search direction for solving an unconstrained minimization problem. Special characters of these directions in the hybrid direction could lead to an improvement of the convergence speed and reduction in the number of function evaluations in the iteration process. Numerical tests on the hybrid directions are performed on the standard test problems (Moré 1981), in particular those with variable dimensions. Comparisons are also made between numerical results obtained from the methods using single directions and hybrid directions. It has been found that the hybrid direction method shows significant reduction in the number of iterations and function evaluations.

KEYWORDS: unconstrained minimization, quasi-Newton, conjugate gradient, steepest descent, hybrid directions

1. INTRODUCTION

Solving unconstrained minimization problems have been continuously developed. The problems are of interest theoretically and importance for applications. Some well-known and classical methods are the steepest descent method, Newton method, conjugate gradient method and the quasi-Newton methods. Some other methods such as optimization bisection (OPTBIS) method for imprecise function and gradient values (Vrahatis 1996) and a dimension reducing (DROPT) optimization method (Grapsa 1996) have also been developed. One common task that most of the methods share is how to obtain the suitable search direction. A widely-used framework for solving unconstrained minimization problems is the line search procedure which requires a descent search direction and a suitably-determined scalar or step length along the search direction in each iteration. The search directions are such as the steepest descent, Newton direction, quasi-Newton, conjugate gradient directions. The criteria for determining a step length are such as the Armijo's rule, backtracking technique (Dennis 1983), Wolfe conditions and strong Wolfe conditions. In the present investigation, the line search framework is used by investigating the performances of various search directions such as the steepest descent, Newton, quasi-Newton and conjugate gradient directions. These directions are also linearly combined to produce a hybrid direction and the line search is performed along this hybrid direction with the mentioned criteria for determining the step length. The ideas of searching for a minimizer along the hybrid direction are motivated by the expanding subspace property based on the conjugate gradient method for minimizing a convex quadratic function, as described in Luenberger, 1984.

3. HYBRID DIRECTION METHOD

Consider the unconstrained minimization problem

$$\min_{x\in R^n} f(x), \tag{3.1}$$

where f is twice continuously differentiable on R^n. The hybrid direction method for solving (3.1) is based on minimization of f over a linear variety $x_c + V_k$, where x_c is the current estimate of a minimizer of f and V_k is the subspace spanned by a set of linearly independent search directions at x_c, $d_0, d_1, \ldots, d_k \in R^n, k < n$. With the line search framework, the new estimate denoted by x_+ is of the form,

$$x_+ = x_c + \lambda v_k, \tag{3.2}$$

where v_k is a vector in V_k and λ is a step length along v_k, determined by the criteria such as Armijo's rule or backtracking technique. Suppose that each of $d_0, d_1, \ldots, d_k$, is a descent direction then so is a linear combination

$$v_k = \alpha_0 d_0 + \alpha_1 d_1 + \ldots + \alpha_k d_k, \tag{3.3}$$

for any positive scalars $\alpha_0, \alpha_1, \ldots, \alpha_k$. The theoretical consideration related to the minimization of f over $x_c + V_k$, is based on the case where f is a convex quadratic function (Sirisathienwatthana 2002). The consideration in the quadratic case is extended to solve (3.1) by a local approximation of f in the form

$$f(x) \approx f(x_c) + \nabla f(x_c)^T (x - x_c) + \frac{1}{2}(x - x_c)^T \nabla^2 f(x_c)(x - x_c). \tag{3.4}$$

The BFGS (Broyden-Fletcher-Goldfarb-Shanno) update is used here to approximate the Hessian in (3.4). The search direction in (3.2) is called the hybrid direction as it is taken from a linear combination of the existing directions. The hybrid directions are of the following forms,

(1) $v = (1-\gamma)d^{PR} + \gamma d^{BFGS}, \quad \gamma = 0, 0.1, \ldots, 1$ (3.5)

(2) $v = \gamma d^{PR} + d^{BFGS}, \quad \gamma = 0, 0.1, \ldots, 1$ (3.6)

(3) $v = d^{SD} + d^{PR} + d^{BFGS}$, and (3.7)

(4) $v = d^{SD} + (1-\gamma)d^{PR} + \gamma d^{BFGS}, \quad \gamma = 0, 0.1, \ldots, 1,$ (3.8)

where d^{SD}, d^{PR} and d^{BFGS} denote the steepest descent, conjugate gradient based on the Polak-Ribière choice of scalar as discussed in Nocedal 1999 and the BFGS quasi-Newton directions, respectively.

Hybrid Direction Algorithm

Given $f: R^n \to R$, $f \in C^2$, a starting point $x_0 \in R^n$, and tol, $\varepsilon > 0$. At iteration j, $j = 0, 1, 2, \ldots$

Step A. Generate the search directions, $d_0^j, d_1^j, \ldots, d_{k-1}^j$.

Step B. Take a linear combination, $v^j = \alpha_0 d_0^j + \alpha_1 d_1^j + \ldots + \alpha_{k-1} d_{k-1}^j$.

Step C. Check the descent property of v^j. If $\nabla f(x_j)^T v^j < 0$ go to Step D., if not, restart with the steepest descent direction, $v^j = -\nabla f(x_j)$.

Step D. Perform the line search from x_j along v^j to obtain the admissible scalar λ_j and set the new estimate as $x_{j+1} = x_j + \lambda_j v^j$.

Step E. Test the admissibility of x_{j+1}. If $\|\nabla f(x_{j+1})\| < \varepsilon$ and $\|x_{j+1} - x_j\| < tol$ stop, else go to Step A.

3. RESULTS AND DISCUSSION

The choices of directions used in the implementation in Step A. are $d_0^j = d_j^{SD}, d_1^j = d_j^{PR}$ and $d_2^j = d_j^{BFGS}$ and the linear combinations used in Step B. are as described in (3.5)-(3.8). The line search routines for determining the step length implemented here for comparison are used here the Wolfe and strong Wolfe conditions as given in Algorithms 3.2 and 3.3 pp. 59-60 (Nocedal 1999). The backtracking techniques are taken from Numerical Recipes in FORTRAN 77: The Art of Scientific Computing (Press 1986-1992). The Armijo's line search is coded as given in Algorithm 1. in Vrahatis, 2000. The test problems are taken from the standard test problems for unconstrained minimization (Moré 1981). The computer codes are in FORTRAN 90 and implemented on a FORTRAN PowerStation 4.0 at the computer laboratory, School of Mathematics, Suranaree University of Technology. Some numerical results of the performance of the hybrid directions (3.5)-(3.8) are shown in Tables 1 – 4, with n, IT and FE denoting the dimension of the test problem, total number of iterations and total number of function evaluations, respectively.

Table 1 shows that as the dimension of the variably dimensioned function gets higher, the hybrid directions $(1-\gamma)d^{PR} + \gamma d^{BFGS}$ with backtracking technique give significant reduction of IT and FE comparing with the performances based on the single directions, d^{SD}, d^{PR} and d^{BFGS}. The hybrid direction $\gamma d^{PR} + d^{BFGS}$ also gives better performances, similarly for $d^{SD} + (1-\gamma)d^{PR} + \gamma d^{BFGS}$. Table 2 also shows the better performance of the hybrid directions for the penalty function I.

Table 1. Results for the Variably Dimensioned Function

Directions	n	γ	Backtracking	Strong Wolfe	Wolfe	Armijo
			IT / FE	IT / FE	IT / FE	IT / FE
SD	32	-	20 / 955	9 / 836	22 / 1116	32 / 1759
PR		-	20 / 955	9 / 836	22 / 1116	32 / 1759
BFGS		-	26 / 918	12 / 1220	13 / 1055	28 / 992
(1) (1-γ)PR+γ BFGS		0.6	9 / 459	9 / 670	13 / 682	38 / 2010
		0.8	12 / 596	9 / 662	13 / 670	38 / 1973
(2) γ PR+ BFGS		0.2	12 / 598	9 / 662	13 / 670	38 / 1973
		0.4	9 / 459	9 / 670	13 / 682	38 / 2010
(3) SD+PR+BFGS		-	13 / 663	9 / 844	22 / 1137	32 / 1790
(4) SD+(1-γ)PR+γ BFGS		0.4	7 / 363	9 / 686	13 / 706	38 / 2084
		0.8	9 / 470	9 / 838	45 / 2246	28 / 1567
SD	64	-	13 / 1154	11 / 1834	27 / 2371	38 / 3495
PR		-	13 / 1154	11 / 1834	27 / 2371	38 / 3495
BFGS		-	32 / 2178	15 / 2901	18 / 2838	44 / 2979
(1) (1-γ)PR+γ BFGS		0.6	11 / 972	12 / 1239	13 / 1189	44 / 3950
		0.8	12 / 1045	12 / 1228	13 / 1177	44 / 3907
(2) γ PR+ BFGS		0.2	13 / 1117	12 / 1228	13 / 1177	44 / 3907
		0.4	11 / 973	12 / 1239	13 / 1189	44 / 3950
(3) SD+PR+BFGS		-	10 / 912	11 / 1894	27 / 2397	38 / 3532
(4) SD+(1-γ)PR+γ BFGS		0.1	11 / 991	9 / 1393	25 / 2260	39 / 3616
		0.2	10 / 886	13 / 1685	21 / 1890	41 / 3784
SD	128	-	21 / 3248	9 / 2949	20 / 3171	41 / 6657
BFGS		-	34 / 4583	17 / 6272	14 / 5110	56 / 7565
(1) (1-γ)PR+γ BFGS		0.2	13 / 2049	11 / 2617	16 / 2632	47 / 7557
(2) γ PR+ BFGS		0.8	12 / 1920	11 / 2617	16 / 2632	47 / 7557
(3) SD+PR+BFGS		-	12 / 1949	9 / 2957	20 / 3190	41 / 6697

Table 2. Results for the Penalty Function I

Directions	n	γ	Backtracking	Strong Wolfe	Wolfe	Armijo
			IT / FE	IT / FE	IT / FE	IT / FE
SD	128	-	diverge	diverge	diverge	diverge
PR		-	diverge	124 / 67087	diverge	diverge
BFGS		-	171/ 22252	116 / 53541	127 / 35438	121 / 15887
(1) $(1-\gamma)$PR+γ BFGS		0.9	38 / 5087	22 / 9998	25 / 7284	55 / 7455
(2) γ PR+ BFGS		0.9	43 / 5793	31 / 11214	45 / 7065	52 / 7150
(3) SD+PR+BFGS		-	44 / 5966	119 / 43626	45 / 7092	58 / 7968
(4) SD+$(1-\gamma)$PR+γ BFGS		0.9	34 / 4595	31 / 11851	57 / 8266	55 / 7548

Table 2 also shows that when the steepest descent or the conjugate direction is implemented alone, divergence occurs; but when it is combined with the BFGS direction, the hybrid directions give m better results in all cases. Tables 3 shows the numerical results for the penalty function II. ' numerical results show the similar situation as in Table 2. That is, when the conjugate directio implemented alone, the divergence occurs; but when it is combined with BFGS direction, the resul hybrid directions (1) and (2) give much better results with the backtracking routine.

Table 3. Results for the Penalty Function II

Directions	n	γ	Backtracking	Strong Wolfe	Wolfe	Armijo
			IT / FE	IT / FE	IT / FE	IT / FE
SD	16	-	diverge	diverge	diverge	diverge
PR	"	-	diverge	158 / 8066	diverge	776 / 17079
BFGS		-	1237 / 21445	248 / 13495	710 / 22227	127 / 2371
(1) $(1-\gamma)$PR+γ BFGS		0.9	281 / 5095	439 / 21137	499 / 10098	214 / 4242
(2) γ PR+ BFGS		0.5	45 / 865	742 / 35730	1455 / 31482	441 / 9513
(3) SD+PR+BFGS		-	2923 / 55565	2383 / 124738	diverge	673 / 15831
(4) SD+$(1-\gamma)$PR+γ BFGS		0.6	diverge	1187 / 59308	diverge	85 / 1933

The hybrid directions are worse than the single direction in the case of the Brown badly sca function. The numerical results obtained here show the promising trend of the hybrid directions speeding up the process in locating a minimizer of the objective function over a linear variety serve as a basis for further theoretical investigation.

4. REFERENCES

Dennis, J.E. Jr., and Schnabel, R.B. (1983). "Numerical methods for unconstrained optimization nonlinear equations." Prentice- Hall, Inc., Englewood Cliffs, NJ.

Grapsa, T.N., and Vrahatis, M.M. (1996). "A dimension-reducing method for unconstrai optimization." *J. Comp. Appl. Math.* Vol. 66, 239-253.

Luenberger, D.G. (1984). "Linear and nonlinear programming." Second edition. Addison-Wes Publishing Company, Reading, MA.

Moré, J.J., Garbow, B.S., and Hillstrom, K.E. (1981). "Testing unconstrained optimization softwa *ACM Trans. Math. Software.* Vol. 7 (1), 17-41.

Nocedal, J., and Wright, S.J. (1999). "Numerical optimization." Springer-Verlag, New York, Inc.,

Sirisathienwatthana, P. (2002). "Line search procedures based on quasi-Newton and conjug gradient directions." M.Sc. Thesis, Suranaree University of Technology, Thailand.

Vrahatis, M.N., Androulakis, G.S., and Manoussakis, G.E.(1996). "A new unconstrained optimization method for imprecise function and gradient values." *J. Math. Anal. Appl.* Vol. 197 (41), 586-607.
Vrahatis, M.N., Androulakis, G.S., Lambrinos, J.N., and Magoulas, G.D. (2000). "A class of gradient unconstrained minimization with adaptive stepsize." *J. Comp. Appl. Math.* Vol.114, 367-386.