DISCUSSION PAPER NO. 206

DUALITY IN UNCONSTRAINED OPTIMIZATION

by

Avinoam Perry

March 1976

Graduate School of Management
Northwestern University
Evanston, Illinois 60201

Duality in unconstrained nonlinear optimization was first introduced by Fletcher [3], who referred to the inverse transformation of the DFP update (Davidon [2], Fletcher and Powell [4]) as a dual update. Oren and Spedicato [8] applied the duality concept to the self-scaling variable metric procedure [6, 7] and showed that dual updates of either the DFP or the BFGS (Broyden [1], Goldfarb [5], Fletcher [3], Shanno [9]) procedures can be obtained directly upon replacing p_k by q_k, S_k by T_k, $\theta = 0$ by $\bar{\theta} = 1$ and vice versa, where

$$p_k \equiv x_{k+1} - x_k$$
$$q_k \equiv \nabla f(x_{k+1}) - \nabla f(x_k) \equiv g_{k+1} - g_k$$
$$S_k = T_k^{-1}$$

More specifically, it was shown that the dual update of

$$S_{k+1} = S_k - \frac{S_k q_k q_k^t S_k}{q_k^t S_k q_k} + \frac{p_k p_k^t}{p_k^t q_k} + \theta v_k v_k^t \tag{1}$$

can be formed by

$$T_{k+1} = T_k - \frac{T_k p_k p_k^t T_k}{p_k^t T_k p_k} + \frac{q_k q_k^t}{q_k^t p_k} + \bar{\theta} w_k w_k^t \tag{2}$$

where θ is either zero or one and $\bar{\theta} = 1 - \theta$

$$v_k = (q_k^t S_k q_k)^{\frac{1}{2}} \left[\frac{p_k}{p_k^t q_k} - \frac{S_k q_k}{q_k^t S_k q_k} \right]$$

$$w_k = (p_k^t T_k p_k)^{\frac{1}{2}} \left[\frac{q_k}{p_k^t q_k} - \frac{T_k p_k}{p_k^t T_k p_k} \right]$$

The purpose of this communication is to show that the existence of dual updates in quasi-Newton algorithms implies the existence of a sequence of equivalent programs which are dual to the original unconstrained quadratic problem

$$\underset{x \in E^n}{\text{minimize}}\; f(x) \tag{3}$$

Definition 1: The quadratic problem

(4) $$\underset{y \in E^n}{\text{minimize}}\ h(y)$$

is a dual of (3) if there exist positive scalars α and ε where $0 < \alpha \leq \varepsilon$ such that the one-dimensional program in α

(5) $$\text{minimize } h(y_k - \alpha T_k \nabla h(y_k))$$

is equivalent to the one-dimensional program in α

(6) $$\text{minimize } f(x_k - \alpha S_k \nabla f(x_k))$$

where T_k and S_k are quasi-Newton updates in the y space and the x space respectively, and T_k is the dual update of S_k.

Theorem 1: Given a quadratic function $f(x)$, $x \in E^n$, there exists a one to one mapping from E^n onto E^n which carries a point $x \in E^n$ onto a point $y \in E^n$ in its dual space such that the conditions specified by Definition 1 are satisfied.

Proof: Denote

(8) $$x = S_k y$$

(9) $$y = T_k x$$

where $T_k = S_k^{-1}$

Then: $f(x) = f(S_k y) = h_k(y)$ and $\nabla h_k(y) = S_k \nabla f(x)$. It follows that

(10) $$y_{k+1} = y_k - \alpha T_k \nabla h_k(y_k) = T_k x_k - \alpha \nabla f(x_k) = T_k x_{k+1}$$

Multiplying both sides of (10) by S_k yields

(11) $$x_{k+1} = x_k - \alpha S_k \nabla f(x_k)$$

which implies the equivalence of (5) and (6).

Another conclusion which follows from (8), (9) is that

(12) $$y_k - y_{k-1} = T_k(x_k - x_{k-1}) = T_k p_{k-1} = q_{k-1}$$

and

(13) $$\nabla h_k(y_k) - \nabla h_k(y_{k-1}) = S_k(\nabla f(x_k) - \nabla f(x_{k-1})) = S_k q_{k-1} = p_{k-1}$$

and by induction we have

(14) $$y_{j+1} - y_j = T_k(x_{j+1} - x_j) = T_k p_j = \nabla f(x_{j+1}) - \nabla f(x_j)$$

(15) $$\nabla h_k(y_{j+1}) - \nabla h_k(y_j) = S_k(\nabla f(x_{j+1}) - \nabla f(x_j)) = S_k q_j = x_{j+1} - x_j$$

for all $j < k$.

This result implies that T_k is a quasi-Newton update of $h_k(y)$. //

It is important to note here that if S_k is a BFGS update in x space then T_k is a DFP update in its dual space.

The economic interpretation of the dual in unconstrained optimization is compatible with the interpretation given to duality in the constrained case. From (12)-(15) it becomes evident that Δy_k can be interpreted as the change in the instantaneous rate of change in f(x) at the new point x_k, while Δx_k is interpreted as the change in the instantaneous rate of change in the dual program. Note that Δx_k is the change in dual of the dual which is identical to the change in the primal variables.

The conclusion of this communication is that the existence of quasi-Newton dual updates implies the existence of a sequence of dual programs in the unconstrained optimization case. This sequence of programs can be generated by applying a sequence of linear transformations which carry $x \in E^n$ onto $y \in E^n$ so that the interpretation of Δy_k is compatible with the common interpretation of duality.

References

[1] Broyden, C. G., "The Convergence of A Class of Double Rank Algorithms," parts I and II, J. Inst. Math. Appl., 7, pp. 76-90, 222-236 (1971).

[2] Davidon, W. C., "Variable Metric Method for Minimization," Argonne National Laboratories Report ANC-5990, rev. (1959).

[3] Fletcher, R., "A New Approach to Variable Metric Algorithms," Computer J., 13, pp. 317-322 (1970).

[4] Fletcher, R. and M.J.D. Powell, "A Rapidly Convergent Descent Method for Minimization," Computer J., 6, pp. 163-168 (1963).

[5] Goldfarb, D., "A Family of Variable Metric Methods Derived By Variational Means," Math. Comp., 24, pp. 23-26 (1970).

[6] Oren, S. S. and D. G. Luenberger, "Self-Scaling Variable Metric Algorithms Part I," Management Science, Vol. 20, No. 5, pp. 845-862 (1974).

[7] Oren, S. S., "Self-Scaling Variable Metric Algorithms Part II," Management Science, Vol. 20, No. 5, pp. 863-874 (1974).

[8] Oren, S. S. and E. Spedicato, "Optimal Conditioning of Self-Scaling Variable Metric Algorithms," presented at ORSA-TIMS Meeting in San Juan (1974).

[9] Shanno, D. F., "Conditioning of Quasi-Newton Methods For Function Minimization," Math. Comp., 24, pp. 647-656 (1970).