The method of conjugate gradients can be generalized to functionals that are not necessarily quadratic. The stepsize parameter ρ_k is still determined by a line search which consists in finding ρ_k such that

$$J(u_k - \rho_k d_k) = \inf_{\rho \in \mathbb{R}} J(u_k - \rho d_k).$$

This is more difficult than in the quadratic case and in general there is no guarantee that ρ_k is unique, so some criterion to pick ρ_k is needed. Then

$$u_{k+1} = u_k - \rho_k d_k,$$

and the next descent direction can be chosen in two ways:

(1) (*Polak–Ribière*)

$$d_k = \nabla J_{u_k} + \frac{\langle \nabla J_{u_k}, \nabla J_{u_k} - \nabla J_{u_{k-1}} \rangle}{\left\| \nabla J_{u_{k-1}} \right\|^2} \, d_{k-1},$$

(2) (*Fletcher–Reeves*)

$$d_k = \nabla J_{u_k} + \frac{\left\| \nabla J_{u_k} \right\|^2}{\left\| \nabla J_{u_{k-1}} \right\|^2} \, d_{k-1}.$$

Consecutive gradients are no longer orthogonal so these methods may run forever. There are various sufficient criteria for convergence. In practice, the Polak–Ribière method converges faster. There is no longer any guarantee that these methods converge to a global minimum.

49.11 Gradient Projection Methods for Constrained Optimization

We now consider the problem of finding the minimum of a convex functional $J \colon V \to \mathbb{R}$ over a nonempty, convex, closed subset U of a Hilbert space V. By Theorem 40.13(3), the functional J has a minimum at $u \in U$ iff

$$dJ_u(v - u) \geq 0 \quad \text{for all } v \in U,$$

which can be expressed as

$$\langle \nabla J_u, v - u \rangle \geq 0 \quad \text{for all } v \in U.$$

On the other hand, by the projection lemma (Proposition 48.5), the condition for a vector $u \in U$ to be the projection of an element $w \in V$ onto U is

$$\langle u - w, v - u \rangle \geq 0 \quad \text{for all } v \in U.$$