where we put a negative sign in front of of the variable ρ_k as a reminder that the descent direction is *opposite* to that of the gradient; a positive value is expected for the scalar ρ_k.

There are four standard methods to pick ρ_k:

(1) *Gradient method with fixed stepsize parameter*. This is the simplest and cheapest method which consists of using the *same* constant $\rho_k = \rho$ for all iterations.

(2) *Gradient method with variable stepsize parameter*. In this method, the parameter ρ_k is adjusted in the course of iterations according to various criteria.

(3) *Gradient method with optimal stepsize parameter*, also called *steepest descent method for the Euclidean norm*. This is a version of Method 2 in which ρ_k is determined by the following line search:

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

This optimization problem only succeeds if the above minimization problem has a *unique* solution.

(4) *Gradient descent method with backtracking line search*. In this method, the step parameter is obtained by performing a backtracking line search.

We have the following useful result about the convergence of the gradient method with optimal parameter.

Proposition 49.13. *Let $J\colon \mathbb{R}^n \to \mathbb{R}$ be an elliptic functional. Then the gradient method with optimal stepsize parameter converges.*

Proof. Since J is elliptic, by Theorem 49.8(3), the functional J has a unique minimum u characterized by $\nabla J_u = 0$. Our goal is to prove that the sequence $(u_k)_{k\geq 0}$ constructed using the gradient method with optimal parameter converges to u, starting from any initial vector u_0. Without loss of generality we may assume that $u_{k+1} \neq u_k$ and $\nabla J_{u_k} \neq 0$ for all k, since otherwise the method converges in a finite number of steps.

Step 1. Show that any two consecutive descent directions are *orthogonal* and

$$J(u_k) - J(u_{k+1}) \geq \frac{\alpha}{2} \left\| u_k - u_{k+1} \right\|^2.$$

Let $\varphi_k\colon \mathbb{R} \to \mathbb{R}$ be the function given by

$$\varphi_k(\rho) = J(u_k - \rho \nabla J_{u_k}).$$

Since the function φ_k is strictly convex and coercive, by Theorem 49.8(2), it has a unique minimum ρ_k which is the unique solution of the equation $\varphi_k'(\rho) = 0$. By the chain rule

$$\begin{aligned}\varphi_k'(\rho) &= dJ_{u_k - \rho\nabla J_{u_k}}(-\nabla J_{u_k}) \\ &= -\langle \nabla J_{u_k - \rho\nabla J_{u_k}}, \nabla J_{u_k} \rangle,\end{aligned}$$