O algoritmo abaixo se refere ao **Método do Gradiente**, para a minimimização de uma função diferenciável $f : \mathbb{R}^n \to \mathbb{R}$. Os parâmetros iniciais, com sugestões de valores, são:

α: taxa de decréscimo suficiente na Regra de Armijo (10^{-4}).

σ: fator de diminuição na busca unidimensional $(0,5)$.

ε: precisão para a norma do gradiente (critério de parada).

M: número máximo de iterações.

x^0: aproximação inicial.

Dados $\alpha \in (0,1)$, $\sigma \in (0,1)$, $\varepsilon > 0$, $M \in \mathbb{N}$ e $x^0 \in \mathbb{R}^n$.

$k = 0$, $g^k = \nabla f(x^k)$, $\eta_k = \|g^k\|_\infty$.

Enquanto $(\eta_k \geq \varepsilon)$ e $(k < M)$ faça:

$\boxed{t_k = 1}$, $w_k = \alpha * (g^k)^T g^k$.

Enquanto $f(x^k - t_k g^k) > f(x^k) - t_k w_k$ faça $t_k \leftarrow \sigma t_k$.

$x^{k+1} = x^k - t_k g^k$.

$g^{k+1} = \nabla f(x^{k+1})$, $\eta_{k+1} = \|g^{k+1}\|_\infty$.

$k \leftarrow k + 1$.

Considere a seguinte modificação para o cálculo do **passo inicial** na busca unidimensional:

$$\boxed{\text{Se } k = 0 \text{ então } t_k = 1 \text{ senão } t_k = \frac{\|x^k - x^{k-1}\|_2}{\|g^k - g^{k-1}\|_2}.}$$

Implemente computacionalmente o **Método do Gradiente** com e sem a modificação acima e compare os desempenhos numéricos para as seguintes funções $f : \mathbb{R}^n \to \mathbb{R}$:

- **Quadrática:** $f(x) = \dfrac{1}{2}\displaystyle\sum_{i=1}^{n} i x_i^2.$

  Mínimo global $f(x^*) = 0$ em $x^* = (0, \ldots, 0)^T$.

- **Rosenbrook:** $f(x) = \displaystyle\sum_{i=1}^{n/2} \left[10(x_{2i} - x_{2i-1}^2)^2 + (x_{2i-1} - 1)^2\right].$

  Mínimo global $f(x^*) = 0$ em $x^* = (1, \ldots, 1)^T$.