Homework assignment #1

1. Barzilai–Borwein step sizes. Consider the gradient method

\[ x_{k+1} = x_k - t_k \nabla f(x_k). \]

We assume \( f \) is convex and differentiable, with \( \text{dom } f = \mathbb{R}^n \), and that \( \nabla f \) is Lipschitz continuous with respect to a norm \( \| \cdot \| \):

\[ \| \nabla f(x) - \nabla f(y) \|_* \leq L \| x - y \| \quad \text{for all } x, y, \]

where \( L \) is a positive constant. Define

\[ s_k = x_k - x_{k-1}, \quad y_k = \nabla f(x_k) - \nabla f(x_{k-1}) \]

and assume \( y_k \neq 0 \). Use the properties in lecture 1 (pages 1.10–1.15) to show that the following two choices for \( t_k \) satisfy \( t_k \geq 1/L \):

\[ t_k = \frac{\|s_k\|^2}{s_k^Ty_k}, \quad t_k = \frac{s_k^Ty_k}{\|y_k\|_*^2}. \]

2. Let \( F(x) = Ax + b \) be an affine function, with \( A \) an \( n \times n \)-matrix. What properties of the matrix \( A \) correspond to the following conditions (a)–(e) on \( F' \)? Distinguish three cases for each subproblem: (1) \( A \) is symmetric, so \( F' \) is the gradient of a quadratic function, (2) \( A \) is skew-symmetric \( (A + A^T = 0) \), and (3) \( A \) is a general non-symmetric matrix.

(a) Monotonicity:

\[ (F(x) - F(y))^T(x - y) \geq 0 \quad \text{for all } x, y. \]

(b) Strict monotonicity:

\[ (F(x) - F(y))^T(x - y) > 0 \quad \text{for all } x \text{ and } y \neq x. \]

(c) Strong monotonicity (for the Euclidean norm):

\[ (F(x) - F(y))^T(x - y) \geq m\|x - y\|_2^2 \quad \text{for all } x, y, \]

where \( m \) is a positive constant.
(d) **Lipschitz continuity (for the Euclidean norm):**

\[ \| F(x) - F(y) \|_2 \leq L \| x - y \|_2 \quad \text{for all } x, y, \]

where \( L \) is a positive constant.

(e) **Co-coercivity (for the Euclidean norm):**

\[ (F(x) - F(y))^T(x - y) \geq \frac{1}{L} \| F(x) - F(y) \|_2^2 \quad \text{for all } x, y, \]

where \( L \) is a positive constant.