HUT , Institute of mathematics Gripenberg
Mat-1.196 Mathematics of neural networks
Exercise 5
12.2–20.2.2002

**1.** Show that if in the conjugate gradient method one chooses the new direction to be

$$\mathbf{s}_{k+1} = -f'(\mathbf{x}_{k+1})^{\mathrm{T}} + \gamma_k \mathbf{s}_k,$$

where

$$\gamma_k = \frac{(f'(\mathbf{x}_{k+1}) - f'(\mathbf{x}_k)) \cdot f'(\mathbf{x}_{k+1})}{f'(\mathbf{x}_k) \cdot f'(\mathbf{x}_k)},$$

then one gets the same sequence of points for quadratic functions as for the standard conjugate gradient method.

**2.** Show that if $0 < \rho \le \sigma < 1$ and $h$ is a continuously differentiable function on $\mathbb{R}$ such that $h$ is bounded from below and $h'(0) < 0$, then there is a number $t > 0$ such that

$$\begin{aligned} h(t) &\le h(0) + t\rho h'(0), \\ |h'(t)| &\le -\sigma h'(0). \end{aligned}$$

**3.** Suppose that $\sigma \in (0, \frac{1}{2})$ and that one in the conjugate gradient method chooses the new point $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{s}_k$ so that

$$\left| f'(\mathbf{x}_{k+1}) \cdot \mathbf{s}_k \right| \le -\sigma f'(\mathbf{x}_k) \cdot \mathbf{s}_k.$$

Show that one then has

$$-\sum_{j=0}^{k} \sigma^j \le \frac{f'(\mathbf{x}_k) \cdot \mathbf{s}_k}{|f'(\mathbf{x}_k)|^2} \le -2 + \sum_{j=0}^{k} \sigma^j.$$

and

$$f'(\mathbf{x}_k) \cdot \mathbf{s}_k < 0,$$

for all $k \ge 0$ unless $f'(\mathbf{x}_k) = 0$.

**C1.** Write a Matlab-function `fun` such that `fun(x, W, dim, sigma`$_1$`,sigma`$_2$`,..)` calculates the the output and intermediate results used by the backpropagation algorithm, when all the weights and thresholds are collected in the vector `W`, the information about the network is in the vector `dim` and the nodefunctions are `sigma`$_1$ etc.