December 22, 2016

HIGH PERFORMANCE SCIENTIFIC COMPUTING — Homework 4

(F) For the first part: Since $k(x) = (\|Ax - b\|_2^2)^2$, by simple chain rule,

$$\nabla k(x) = 2\|Ax - b\|_2^2 \nabla_x (\|Ax - b\|_2^2) = 2\|Ax - b\|_2^2 (2A^T(Ax - b)) = 4\|Ax - b\|_2^2 A^T(Ax - b).$$

The dimensions of the gradient vector $\nabla k(x)$ is $n \times 1$. (Recall that for any real-valued differentiable function g with domain $\mathbb{R}^\ell$, the gradient at any point would be an ℓ-vector.)

For the second part: Note that since v is the vector of all ones,

$$h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_n(x) \end{bmatrix}, \text{ where } h_i(x) = x^T M x + x_i, \text{ for } i = 1, \ldots, n.$$

Noting that for $i = 1, \ldots, n$,

$$\nabla h_i(x) = 2Mx + e_i,$$

where e_i is the i-th column of the $n \times n$ identity matrix I_n, we have

$$Jh(x) = \begin{bmatrix} \nabla h_1(x)^T \\ \vdots \\ \nabla h_n(x)^T \end{bmatrix} = \begin{bmatrix} 2x^T M + e_1 \\ \vdots \\ 2x^T M + e_n \end{bmatrix} = \begin{bmatrix} 2x^T M \\ \vdots \\ 2x^T M \end{bmatrix} + I_n,$$

which can also be conveniently expressed as $2vx^T M + I_n$.