Quiz 1

1. [3 pts] Remember that for the trace operation $tr(v \cdot v^\top) = \sum_i v_i^2 = v^\top v$, and if a square matrix P is orthonormal then $P^\top P = I$ (the identity). Given that $v = [1 \quad 0 \quad 2]^\top$, and R is an orthonormal matrix, what is $tr(Rvv^\top R^\top)$ equal to?

2. [3 pts] Given that the A_{12}, A_{21} and $A_{22} \in \mathbb{R}^{3\times3}$ are invertible, and $b_1, b_2, x_1, x_2 \in \mathbb{R}^3$, then solve for x_1 and x_2 from

$$\begin{bmatrix} 0 & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

3. [4 pts] Consider the function $f(x) = \sum_{i=1}^N a_i \cdot \ln(x_i)$ where $a = [a_1 \quad \ldots \quad a_N]^\top$ is a constant vector, and the function $\rho(x)$ which is the sigmoid function applied element-wise to x, i.e., $\rho(x) = [\sigma(x_1) \quad \cdots \quad \sigma(x_N)]^\top$, where $\sigma(u) = \frac{1}{1+e^{-u}}$. Derive an expression for $\nabla_x f\big(\rho(x)\big)$ in vector form (i.e., without summations).

   Hint: Remember that $\sigma'(u) = \sigma(u) \cdot (1 - \sigma(u))$.

Answer

1. By definition the expression is equal to $(Rv)^\top(Rv) = \ v^\top R^\top Rv = v^\top v = 5$
2. From the first set of questions $x_2 = A_{12}^{-1}b_1$, by substituting into the second equation we get $x_1 = A_{21}^{-1}\left(b_2 - A_{22}A_{12}^{-1}b_1\right)$.
3. Let us define $F(x) = f(\rho(x))$. By chain rule we have
$$\frac{\partial F}{\partial x_k}(x) = \sum_i a_i \cdot \left(\frac{1}{\sigma(x_i)}\right) \cdot \sigma(x_i) \cdot \left(1 - \sigma(x_i)\right) \cdot \frac{\partial x_i}{\partial x_k} = a_k \cdot (1 - \sigma(x_k))$$

So,
$$\nabla_x F(x) = A \cdot \left(1 - \rho(x)\right)$$
where $A = diag(a)$.