Numerical Linear Algebra and Approximation: Solutions for Sheet 1

1. $H(w)$ is a Householder matrix (often called a 'Householder reflection'), which has the form $H(w) := I - 2ww^{\mathrm{T}}/w^{\mathrm{T}}w$. Compute
$$H(w)H(w) = \left(I - \frac{2ww^{\mathrm{T}}}{w^{\mathrm{T}}w}\right)\left(I - \frac{2ww^{\mathrm{T}}}{w^{\mathrm{T}}w}\right) = I - 4\frac{ww^{\mathrm{T}}}{w^{\mathrm{T}}w} + 4\frac{w(w^{\mathrm{T}}w)w^{\mathrm{T}}}{(w^{\mathrm{T}}w)(w^{\mathrm{T}}w)} = I - 4\frac{ww^{\mathrm{T}}}{w^{\mathrm{T}}w} + 4\frac{ww^{\mathrm{T}}}{w^{\mathrm{T}}w} = I.$$
Note that H is both orthogonal and symmetric: $H = H^{\mathrm{T}} = H^{-1}$.

2. Verify the three properties required satisfied by a vector norm:
1) positivity: $\max_i |x_i| \geq 0, \max_i |x_i| = 0 \iff x = 0$ by properties of $|\cdot|$;
2) scaling: $|\alpha x_i| = |\alpha||x_i| \implies \max_i |\alpha x_i| = |\alpha| \max_i |x_i|$;
3) triangle inequality: $|x_i + y_i| \leq |x_i| + |y_i| \implies \max_i |x_i + y_i| \leq \max_i |x_i| + \max_i |y_i|$.
Thus $\|\cdot\|_\infty$ is a vector norm.

3. Suppose $\|\cdot\|$ is some vector norm. Verify the three requirements for matrix norms:
1) positivity: $\|x\| \geq 0\,\forall x \implies \|Ax\| \geq 0\,\forall x$, so $\|Ax\|/\|x\| \geq 0$ for $x \neq 0$;
If $\|A\| = 0$ then $\|Ax\| = 0\,\forall x \implies Ax = 0\,\forall x \implies A = 0$. Clearly $\|0\| = 0$.
2) scaling: $\|\alpha x\| = |\alpha|\,\|x\| \implies \|(\alpha A)x\| = |\alpha|\,\|Ax\|$, so $\|(\alpha A)x\|/\|x\| = |\alpha|\,\|Ax\|/\|x\|\,\forall x$;
3) triangle inequality: $\|x+y\| \leq \|x\| + \|y\| \implies \|A(x+y)\| \leq \|Ax\| + \|Ay\|$,
so $\sup_x \|(A+B)x\|/\|x\| \leq \sup_x \|Ax\|/\|x\| + \sup_x \|Bx\|/\|x\|$.
To show $\|AB\| \leq \|A\|\,\|B\|$, consider
$$\begin{aligned}\|AB\| = \sup_x \frac{\|ABx\|}{\|x\|} = \sup_x \frac{\|ABx\|}{\|x\|}\frac{\|Bx\|}{\|Bx\|} &\leq \sup_x \left(\frac{\|ABx\|}{\|Bx\|}\right) \sup_x \left(\frac{\|Bx\|}{\|x\|}\right) \\ &\leq \sup_y \left(\frac{\|Ay\|}{\|y\|}\right) \sup_x \left(\frac{\|Bx\|}{\|x\|}\right) = \|A\|\,\|B\|.\end{aligned}$$

4. Let $A \in \mathbb{R}^{m\times n}$. From the operator norm definition, we have $\|A\|_1 = \sup_{\|x\|_1=1} \|Ax\|_1$. Our strategy will be to develop an upper bound for $\|A\|_1$, and then demonstrate that the bound is attained for a specific x. For general $x \in \mathbb{R}^n$ with $\|x\|_1 = 1$, we have
$$\begin{aligned}\|Ax\|_1 &= \sum_{i=1}^m \left|\sum_{j=1}^n \alpha_{ij} x_j\right| \\ &\leq \sum_{i=1}^m \sum_{j=1}^n |\alpha_{ij}|\,|x_j| = \sum_{j=1}^n \left(|x_j| \sum_{i=1}^m |\alpha_{ij}|\right) \\ &\leq \left(\max_j \sum_{i=1}^m |\alpha_{ij}|\right) \sum_{j=1}^n |x_j| \\ &= \max_j \sum_{i=1}^m |\alpha_{ij}|.\end{aligned}$$
This bound is actually attained by the vector which is all zeros except for a 1 in the entry corresponding to the column of A with maximum 1-norm. Thus, $\|A\|_1 = \max_j \sum_{i=1}^m |\alpha_{ij}|$.

5. Suppose $A \in \mathbb{R}^{m\times n}$ and let $Q \in \mathbb{R}^{m\times m}$ be orthogonal. Recall that $\|Qx\|_2 = \|x\|_2$ since $\|Qx\|_2^2 = (Qx)^{\mathrm{T}}(Qx) = x^{\mathrm{T}}Q^{\mathrm{T}}Qx = x^{\mathrm{T}}x = \|x\|_2^2$.
Partition $A \in \mathbb{R}^{m\times n}$ by columns, $A = [a_1\ a_2\ \cdots\ a_n]$, where $a_j \in \mathbb{R}^m$. Then we can write the Frobenius norm of A using dot products:

$$\|A\|_{\mathrm{F}}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} |\alpha_{ij}|^2 = \sum_{j=1}^{n} a_j^{\mathrm{T}} a_j.$$

Now write $QA = Q[a_1\ a_2\ \cdots\ a_n] = [Qa_1\ Qa_2\ \cdots\ Qa_n]$, and compute

$$\|QA\|_{\mathrm{F}}^2 = \sum_{j=1}^{n}(Qa_j)^{\mathrm{T}}(Qa_j) = \sum_{j=1}^{n} a_j^{\mathrm{T}}Q^{\mathrm{T}}Qa_j = \sum_{j=1}^{n} a_j^{\mathrm{T}}a_j = \|A\|_{\mathrm{F}}^2.$$

6. Let $A = U\Sigma V^{\mathrm{T}}$ be the singular value decomposition (SVD) for A. (Recall that U and V are orthogonal, and $\Sigma = \mathrm{diag}(\sigma_1,\ldots,\sigma_n)$ is diagonal with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.) Then

$$\|A\|_2 = \sup_{\|x\|=1}\|Ax\|_2 = \sup_{\|x\|=1}\|U\Sigma V^{\mathrm{T}}x\| = \sup_{\|x\|=1}\|\Sigma V^{\mathrm{T}}x\| = \sup_{\|V^{\mathrm{T}}x\|=1}\|\Sigma V^{\mathrm{T}}x\| = \sup_{\|y\|=1}\|\Sigma y\|.$$

This supremum is attained by the unit vector $y = [1,0,\ldots,0]^{\mathrm{T}}$, with $\|A\|_2 = \sigma_1$.

7. Let $A \in \mathbb{R}^{m\times n}$. The SVD of A is $A = U\Sigma V^{\mathrm{T}}$. Suppose $m \geq n$. Then $A^{\mathrm{T}}A = V\Sigma^{\mathrm{T}}U^{\mathrm{T}}U\Sigma V^{\mathrm{T}} = V\Sigma^{\mathrm{T}}\Sigma V^{\mathrm{T}}$. Notice that $\Sigma^{\mathrm{T}}\Sigma = \mathrm{diag}(\sigma_1^2,\sigma_2^2,\ldots,\sigma_n^2)$. Since $\Sigma^{\mathrm{T}}\Sigma$ is diagonal and $V^{\mathrm{T}} = V^{-1}$, this is an eigenvalue-eigenvector decomposition of $A^{\mathrm{T}}A$. Thus, the eigenvalues of $A^{\mathrm{T}}A$ are $\sigma_1^2,\sigma_2^2,\ldots,\sigma_n^2$.
Note that $A^{\mathrm{T}}A$ is square ($n\times n$), symmetric, and has positive eigenvalues.
The proof for $m \leq n$ is very similar.

For the MATLAB exercise, you should have found that `svd(A)` returns the values 9.5255 and 0.5143 (rounding to four digits after the decimal point), while `eig(A'*A)` returns the values 0.2645 and 90.7355 (also rouding to four digits). Observe that these agree with our expectations from the first part of the problem.

8. Suppose $A \in \mathbb{R}^{n\times n}$ is nonsingular with the SVD $A = U\Sigma V^{\mathrm{T}}$. Then we seek a matrix B with $AB = I$. We can compute B: $U\Sigma V^{\mathrm{T}}B = I \implies B = V\Sigma^{-1}U^{\mathrm{T}}$. Note that $\Sigma^{-1} = \mathrm{diag}(1/\sigma_1, 1/\sigma_2,\ldots,1/\sigma_n)$. Although $A^{-1} = B = V\Sigma^{-1}U^{\mathrm{T}}$, this is not a proper SVD because it violates our convention for ordering the entries of Σ. To fix this problem, we need the permutation matrix with ones in the $(i, n-i+1)$ entries,

$$P = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{pmatrix}.$$

Unspecified entries are zero. Note that this matrix is orthogonal, $PP^{\mathrm{T}} = I$.

Then we have $P\Sigma^{-1}P^{\mathrm{T}} = \mathrm{diag}(1/\sigma_n,\ldots,1/\sigma_2,1/\sigma_1)$, which obeys the desired ordering convention. Thus, $A^{-1} = VP^{\mathrm{T}}(P\Sigma^{-1}P^{\mathrm{T}})PU^{\mathrm{T}}$ is the SVD of A^{-1}.