1. Suppose $W$ is an $m \times n$ matrix with orthonormal columns. What is $\|W\|_F$?

$$\|W\|_F = \sqrt{n}$$

2. Check carefully that if $O$ is an orthonormal $m \times m$ matrix, and if $A$ is an $m \times n$ matrix, $\|OA\|_F = \|A\|_F$.

Since the $j$th column of $OA$, $(OA)_j = OA_j$,

$$\|OA\|_F^2 = \sum_{j=1}^n \|OA_j\|^2 = \sum_{j=1}^n \|A_j\|^2 = \|A\|_F^2$$

3. If $A = USV^T$ is an SVD of $A$, write a formula for $\|A\|_F$ in terms of the diagonal elements of $S$.

$$\|A\|_F = \|U^TAV\|_F = \|S\|_F = \sqrt{\sum_{j=1}^{\min(m,n)} S_{jj}^2}$$

4. Suppose $W$ is an $m \times n$ matrix with orthonormal columns. Show that the norm of each row of $W$ is less than or equal to 1.

We use the fact that the transpose of a square orthonormal matrix is also orthonormal (because in that case, the transpose is the inverse, and the inverse of the transpose is the transpose of the inverse). Since the columns of $W$ are orthonormal, $n \leq m$. If $n = m$, then $W$ has orthonormal rows as well, and the statement follows. If $n < m$, let $V$ be an orthonormal basis for $\text{col}(W)^\perp$, and let

$$Q = [W \ V].$$

Then $Q$ is an $m \times m$ orthonormal matrix, so the rows of $Q$ have norm 1, and so for each $i$

$$\sum_{j=1}^n W_{ij}^2 = \sum_{j=1}^n Q_{ij}^2 \leq \sum_{j=1}^m Q_{ij}^2 = 1.$$ 

5. Work through the manipulations in equations (4.13), (4.14), and (4.15).

See the notes on least squares; remember that if the $m \times r$ matrix $W$ has orthonormal columns, $W^TW = I_r$, and that this means $\|W\alpha\| = ||\alpha||$ for any $\alpha \in \mathbb{R}^r$.

6. Find the flaw in the following argument going straight from (4.2) to the conclusion: Since $S$ is zero for non-diagonal entries, it is clear that you pay a penalty for any off diagonal entry in $C$ in (??). Thus we may conclude that $C$ is also diagonal. The only diagonal matrices of rank $r$ are nonzero in precisely $r$ locations on the diagonal; since $S_{jj}$ is non-decreasing in $j$, keeping the first $r$ entries gives the least error.

Because of the low rank constraint on $C$, changing the diagonal entries can affect the off diagonal entries, and vice versa. So it may be that paying a penalty on the off diagonal entries may result in a lower penalty on the diagonal entries, and so it is not possible to directly conclude that $C$ is diagonal.