Question 1

It is a well known theorem that all possible solutions to the LS problem comply with the normal equation:

$$X^T X w = X^T y$$

Using the X's SVD form, and some simplification:

$$V \Sigma^T U^T U \Sigma V^T w = V \Sigma^T U^T y$$

$$V \Sigma^T \Sigma V^T w = V \Sigma^T U^T y$$

$$\Sigma^T \Sigma V^T w = \Sigma^T U^T y$$

Denoting $z = V^T w$ and $\hat{y} = U^T y$:

$$\Sigma^T \Sigma z = \Sigma^T U^T \hat{y}$$

Let's look at this matrix quation:

$$\begin{pmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_r^2 & \\ & & & & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_r \\ z_r + 1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_r & \\ & & & & 0 \end{pmatrix} \begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_m \end{pmatrix}$$

Thus a solution can be written as:

$$z = \begin{pmatrix} \frac{\hat{y}_1}{\sigma_1} \\ \vdots \\ \frac{\hat{y}_r}{\sigma_r} \\ z_r + 1 \\ \vdots \\ z_n \end{pmatrix}$$

Where we have freedom of choice for the z_{r+1} to z_n values.
We should regard that this solution is not gurenteed to be optimal.

It's interesting to see that if we choose z_{r+1} to z_n to be 0, we get the most optimal solution in terms of L_2 norm.
This optimal solution means that $w = Vz$ is also an optimal solution (because V is orthonormal, and thus doesn't affect the norm).
Reconstructing this solution:

$$\hat{w} = V \begin{pmatrix} \sigma_1^{-1} & & & & \\ & \sigma_2^{-1} & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_r^{-1} & \\ & & & & 0 \end{pmatrix} U^T y$$

This $\hat{w}$ solution is also our solution using the defintion of Σ^+:

$$(\Sigma^+)^2 \Sigma^T = \begin{pmatrix} \sigma_1^{-2} & & & & \\ & \sigma_2^{-2} & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_r^{-2} & \\ & & & & 0 \end{pmatrix} \begin{pmatrix} \sigma_1 & & & & \\ & \sigma_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_r & \\ & & & & 0 \end{pmatrix}$$

$$(\Sigma^+)^2 \Sigma^T = \begin{pmatrix} \sigma_1^{-1} & & & & \\ & \sigma_2^{-1} & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_r^{-1} & \\ & & & & 0 \end{pmatrix}$$

$$\hat{w} = V (\Sigma^+)^2 \Sigma^T U^T y$$