Assignment 1

[illegible]

1.) A.) Assume $m_\alpha = (K^T K + \alpha I)^{-1} K^T d$

Let U and V be the matrices containing the singular vectors of K, and let S be the diagonal matrix containing the singular values of K, then:

$$m_\alpha = \left((USV^T)^T (USV^T) + \alpha I\right)^{-1} (USV^T)^T d$$

$$= (VS^T \underbrace{U^T U}_{I} S V^T + \alpha I)^{-1} V S^T U^T d$$

$$= (VS^T S V^T + \alpha I)^{-1} V S^T U^T d = \frac{\cancel{V} S^T}{(\cancel{V} S^T S V^T + \alpha I)} U^T d$$

V^{-1}

$$= \frac{S^T}{(S^T S V^T + V^{-1} \alpha I)} U^T d = \frac{S^T}{V^{-1}(S^T S + \alpha I)} U^T d$$

$V^T = V^{-1}$ $S^T = S$

$$= V \frac{S}{(S^T S + \alpha I)} U^T d = \boxed{\sum_{i=1}^{N} \frac{s_i}{s_i^2 + \alpha} (u_i^T d)\, v_i}$$

For $K^T K + \alpha I$ we have:

$$(USV^T)^T (USV) + \alpha I = VS^T U^T U S V^T + \alpha I$$

$$= VS^T S V^T + \alpha I$$

$VS^T S V^T$ is the eigen decomposition of the matrix $K^T K$, since s_N is strictly positive, this implies that $K^T K$ is positive definite. αI is also positive definite by construction since $\alpha > 0$

$$\boxed{\text{Therefore } K^T K + \alpha I \text{ is positive definite, and also invertible}}$$

B.) $\min_m \frac{1}{2}\|Km - d\|^2 + \frac{\alpha}{2}\|m\|^2$

$$J(m) = \frac{1}{2}(Km-d)^T(Km-d) + \frac{\alpha}{2} m^T m$$

We consider:

$\varepsilon \in \mathbb{R}$, $\tilde{m} \in \mathbb{R}^n$

$$f(\varepsilon) = \frac{1}{2}(K(m+\varepsilon\tilde{m})-d)^T(K(m+\varepsilon\tilde{m})-d) + \frac{\alpha}{2}(m+\varepsilon\tilde{m})^T(m+\varepsilon\tilde{m})$$

$$= \frac{1}{2}\left[(K(m+\varepsilon\tilde{m}))^T K(m+\varepsilon\tilde{m}) - d^T K(m+\varepsilon\tilde{m}) - (K(m+\varepsilon\tilde{m}))^T d + d^T d\right] + \frac{\alpha}{2}(m+\varepsilon\tilde{m})^T(m+\varepsilon\tilde{m})$$

$$= \frac{1}{2}(m+\varepsilon\tilde{m})^T K^T K(m+\varepsilon\tilde{m}) \underbrace{- \frac{1}{2} d^T K(m+\varepsilon\tilde{m}) - \frac{1}{2}(m+\varepsilon\tilde{m})^T K^T d} + \frac{1}{2}d^T d + \frac{\alpha}{2}(m+\varepsilon\tilde{m})^T(m+\varepsilon\tilde{m})$$

$$= \frac{1}{2}(m+\varepsilon\tilde{m})^T K^T K(m+\varepsilon\tilde{m}) - (m+\varepsilon\tilde{m})^T K^T d + \frac{1}{2}d^T d + \frac{\alpha}{2}(m+\varepsilon\tilde{m})^T(m+\varepsilon\tilde{m})$$

$$= \frac{1}{2}\left[m^T K^T K m + \underbrace{m^T K^T K \varepsilon\tilde{m} + \varepsilon\tilde{m}^T K^T K m} + \varepsilon\tilde{m}^T K^T K \varepsilon\tilde{m}\right] - m^T K^T d - \varepsilon\tilde{m}^T K^T d + \frac{1}{2}d^T d$$

$$+ \frac{\alpha}{2}\left[m^T m + \underbrace{m^T\varepsilon\tilde{m} + \varepsilon\tilde{m}^T m} + \varepsilon\tilde{m}^T\varepsilon\tilde{m}\right]$$

$$= \frac{1}{2}m^T K^T K m + \varepsilon\tilde{m}^T K^T K m + \frac{1}{2}\varepsilon^2\tilde{m}^T K^T K\tilde{m} - m^T K^T d - \varepsilon\tilde{m}^T K^T d + \frac{1}{2}d^T d + \frac{\alpha}{2}m^T m$$

$$+ \alpha\varepsilon\tilde{m}^T m + \frac{\alpha}{2}\varepsilon^2\tilde{m}^T\tilde{m}$$

$$\frac{df(\varepsilon)}{d\varepsilon} = \frac{dJ(m+\varepsilon e_i)}{d\varepsilon} = \frac{dJ}{dm_i} \quad \text{where } \tilde{m} = e_i \quad \tilde{m} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \leftarrow \text{ith element}$$

$$= 0 + \tilde{m}^T K^T K m + \varepsilon\tilde{m}^T K^T K\tilde{m} - 0 - \tilde{m}^T K^T d + 0 + 0 + \alpha\tilde{m}^T m + \alpha\varepsilon\tilde{m}^T\tilde{m}$$

$$\left.\frac{df(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} = \tilde{m}^T K^T K m + 0 - \tilde{m}^T K^T d + \alpha\tilde{m}^T m + 0 = \tilde{m}^T(K^T K m - K^T d + \alpha m)$$

For one element of m

Add for all elements

$$\frac{dJ}{dm} = K^T K m - K^T d + \alpha I m$$

Now we find min of $J(m)$

$$\frac{dJ}{dm} = 0 = K^T K m_\alpha - K^T d + \alpha m_\alpha$$

$$\Rightarrow K^T K m_\alpha + \alpha I m_\alpha = K^T d$$

$$(K^T K + \alpha I) m_\alpha = K^T d$$

$$\boxed{m_\alpha = (K^T K + \alpha I)^{-1} K^T d}$$