Assignment 2b, Numerical Optimization & Large Scale Linear Algebra
Professor: P. Vassalos

Stratos Gounidellis, DS3517005

Problem 1

$$\min_f \left\| \begin{bmatrix} g \\ 0 \end{bmatrix} - \begin{bmatrix} K \\ \alpha I \end{bmatrix} f \right\|_2^2 = \min_f \left\| \begin{bmatrix} g - Kf \\ -\alpha I f \end{bmatrix} \right\|_2^2 \quad (1)$$

It holds that the squared l_2-norm of a vector is equivalent to the sum of the squares of its components. More specifically, given a vector $x = [x_1, x_2, \dots, x_m]$ we get:

$$\begin{aligned} \|x\|_2^2 &= x_1^2 + x_2^2 + x_3^2 + \dots + x_m^2 + x_{m+1}^2 + \dots + x_n^2 \\ &= (x_1^2 + x_2^2 + \dots x_m^2) + (x_{m+1}^2 + x_{m+2}^2 + x_{m+3}^2 \dots + x_n^2) \\ &= \|x_1^m\|_2^2 + \|x_{m+1}^n\|_2^2 \end{aligned}$$

Consequently, from the above relationships:

$$\begin{aligned} \min_f \left\| \begin{bmatrix} g - Kf \\ -\alpha I f \end{bmatrix} \right\|_2^2 &= \min_f \{ \|g - Kf\|_2^2 + \|-\alpha I f\|_2^2 \} = \\ &= \min_f \{ \|g - Kf\|_2^2 + \alpha^2 \|f\|_2^2 \} \quad (2) \end{aligned}$$

It is proved that relationships (1) and (2) are equivalent.

Problem 2

$$\min_f \{ \|g - Kf\|_2^2 + \alpha^2 \|f\|_2^2 \}$$

⇝ We will work in each component separately

- $\|g - Kf\|_2^2 = \|U^T g - U^T K f\|_2^2 = \|U^T g - U^T U \Sigma V^T f\|_2^2$
$$= \|U^T g - \Sigma V^T f\|_2^2$$

- $\alpha^2 \|f\|_2^2 = \alpha^2 \|V^T f\|_2^2$

We set $\hat{g} = U^T g$ and $\hat{f} = V^T f$ so we get

$\|g - Kf\|_2^2 = \|\hat{g} - \Sigma \hat{f}\|_2^2$ and $\alpha^2 \|f\|_2^2 = \alpha^2 \|\hat{f}\|_2^2$ respectively

The above matrix norm multiplications were implemented based on the property that when we multiply a norm with an orthogonal matrix, the norm remains unchanged.

So, $\min_f \{ \|g - Kf\|_2^2 + \alpha^2 \|f\|_2^2 \} = \min_f \{ \|\hat{g} - \Sigma \hat{f}\|_2^2 + \alpha^2 \|\hat{f}\|_2^2 \}$ (1)

(1) ⇒ $\min_f \left\| \begin{bmatrix} \hat{g} \\ 0 \end{bmatrix} - \begin{bmatrix} \Sigma \\ \alpha I \end{bmatrix} \hat{f} \right\|_2^2$

Problem 3

Let $L(\hat{f}) = \|\hat{g} - \Sigma \hat{f}\|_2^2 + a^2 \|\hat{f}\|_2^2$.

Then taking the derivative of $L(\hat{f})$ we get:

$$\frac{\partial L(\hat{f})}{\partial \hat{f}} = -2\Sigma^T(\hat{g} - \Sigma \hat{f}) + 2a^2 \hat{f}$$

Set the derivative equal to zero and solve for $\hat{f}$

$$-2\Sigma^T(\hat{g} - \Sigma\hat{f}) + 2a^2\hat{f} = 0 \Rightarrow -2\Sigma^T\hat{g} + 2\Sigma^T\Sigma\hat{f} + 2a^2\hat{f} = 0 \Rightarrow$$

$$-\Sigma^T\hat{g} + \Sigma^T\Sigma\hat{f} + a^2 f = 0 \Rightarrow (\Sigma^T\Sigma + a^2 I)\hat{f} = \Sigma^T\hat{g} \Rightarrow$$

$$\hat{f} = \cancel{V} (\cancel{\Sigma^T} (\Sigma^T\Sigma + a^2 I)^{-1} \Sigma^T \hat{g} \Rightarrow V^T f = (\Sigma^T\Sigma + a^2 I)^{-1} \Sigma^T U^T g$$

$$\Rightarrow f = V(\Sigma^T\Sigma + a^2 I)^{-1}\Sigma^T U^T g \Rightarrow \boxed{f = \sum_{i=1}^{\mathrm{rank}(K)} \sigma_i \frac{u_i^T g\, v_i}{\sigma_i^2 + a^2}},$$

where u_i is the ith column of U and v_i is the ith column of V.

Problem 4

Let $L(f) = \|g - Kf\|_2^2$

Then taking the derivative of $L(f)$ wrt f we get:

$$\frac{\partial L(f)}{\partial f} = -2K^T(g - Kf)$$

Set the derivative equal to zere and solve for f:

$$-2K^T(g-Kf) = 0 \Rightarrow K^T g - K^T K f = 0 \Rightarrow -V\Sigma^T U^T g + V\Sigma^T U^T U \Sigma V^T f = 0$$

$$\Rightarrow (V\Sigma^T \Sigma V^T) f = V\Sigma^T U^T g \Rightarrow f = (V\Sigma^T \Sigma V^T)^{-1} V\Sigma^T U^T g \Rightarrow (*)$$

$$f = (V\Sigma^2 V^T)^{-1} V\Sigma^T U^T g \Rightarrow f = V\Sigma^{-2} \Sigma U^T g \Rightarrow f = V\Sigma^{-1} U^T g$$

$$\Rightarrow f = \sum_{i=1}^{\mathrm{rank}(K)} \frac{u_i^T g}{\sigma_i} v_i$$

(*) Matrix S in our case is square diagonal with entries σ_1, σ_2, .. σ_n. So, the S^{-1} will have the elements $\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \frac{1}{\sigma_n}$ in its main diagonal and zeros elsewhere.