Bruce James

Assignment 4:

1.
(a) We start by plugging $\mathbf{c} - \mathbf{x}(\mathbf{c}^T\mathbf{x})$ in for $\mathbf{y}$ to find

$$\min_{\mathbf{x},\mathbf{y}} \left\{\sum_{1\le i\le n} |\mathbf{a}_i - \mathbf{y}|^2 - |\mathbf{x}^T\mathbf{a}_i|^2, \text{ subject to } \mathbf{y}^T\mathbf{x} = 0,\ \mathbf{x}^T\mathbf{x} = 1\right\}, \qquad (1)$$

and simplifying the expression $|\mathbf{a}_i - \mathbf{y}|^2 - |\mathbf{x}^T\mathbf{a}_i|$, as such:

$$\begin{aligned}
|\mathbf{a}_i - \mathbf{y}|^2 - |\mathbf{x}^T\mathbf{a}_i| &= [(\mathbf{a}_i - \mathbf{c}) + \mathbf{x}(\mathbf{c}^T\mathbf{x})]^T[(\mathbf{a}_i - \mathbf{c}) + \mathbf{x}(\mathbf{c}^T\mathbf{x})] - |\mathbf{x}^T\mathbf{a}_i|^2 \\
&= |\mathbf{a}_i - \mathbf{c}|^2 + 2(\mathbf{a}_i - \mathbf{c})^T\, \mathbf{x}(\mathbf{c}^T\mathbf{x}) \\
&\qquad + ((\mathbf{c}^T\mathbf{x})^2\,)\mathbf{x}^T\mathbf{x} - (\mathbf{x}^T\mathbf{a}_i)^T(\mathbf{x}^T\mathbf{a}_i) \\
&= |\mathbf{a}_i - \mathbf{c}|^2 - [-2(\mathbf{a}_i - \mathbf{c})^T\, \mathbf{x}(\mathbf{c}^T\mathbf{x}) \\
&\qquad - ((\mathbf{c}^T\mathbf{x})^2\,)\mathbf{x}^T\mathbf{x} + (\mathbf{x}^T\mathbf{a}_i)^T(\mathbf{x}^T\mathbf{a}_i)] \\
&= |\mathbf{a}_i - \mathbf{c}|^2 - [-2\mathbf{a}_i^T\mathbf{x}(\mathbf{c}^T\mathbf{c}) + 2\mathbf{c}^T\mathbf{x}(\mathbf{c}^T\mathbf{c}) \\
&\qquad - (\mathbf{c}^T\mathbf{x})^2 + (\mathbf{x}^T\mathbf{a}_i)^T(\mathbf{x}^T\mathbf{a}_i)] \\
&= |\mathbf{a}_i - \mathbf{c}|^2 - [-2\mathbf{a}_i^T\mathbf{x}(\mathbf{c}^T\mathbf{c}) + (\mathbf{c}^T\mathbf{x})^2 + (\mathbf{x}^T\mathbf{a}_i)^2] \\
&= |\mathbf{a}_i - \mathbf{c}|^2 - |\mathbf{c}^T\mathbf{x} - \mathbf{a}_i^T\mathbf{x}|^2.
\end{aligned}$$

Since $|\mathbf{a}_i - \mathbf{c}|^2 \ge 0$, and the only expression that depends on $\mathbf{x}$ is

$$|\mathbf{c}^T\mathbf{x} - \mathbf{a}_i^T\mathbf{x}|^2 \ge 0,\ \forall\ \mathbf{x} \in \mathbf{R}^n,$$

then the minimum of (1) is the largest value of $(\mathbf{c}^T\mathbf{x} - \mathbf{a}_i^T\mathbf{x})^2$, namely

$$\sup_{\mathbf{x}} \left\{\sum_{1\le i\le n} |\mathbf{c}^T\mathbf{x} - \mathbf{a}_i^T\mathbf{x}|^2, \text{ constrained to } \mathbf{x}^T\mathbf{x} = 1\right\}. \qquad (2)$$

In order to set-up the optimization problem, let $f(\mathbf{x})$ be the vector-valued objective function

$$f(x_1, \ldots, x_n) = \sum_{1\le i\le n} \left[|\mathbf{c}^T\mathbf{x} - \mathbf{a}_i^T\mathbf{x}|^2\right],$$

and

$$g(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$$

be the constraint function. Taking the gradient of f yields

$$\nabla f(\mathbf{x}) = \sum_{1\le j\le n} f_j(x_1, \ldots, x_n)\ \mathbf{e}_j,$$

where f_j is the partial derivative of f with respect to the j-th component, and $\mathbf{e}_j$ is the unit basis vector of the j-th component namely $\mathbf{e}_j=[0,\ldots,e_j=1,\ldots,0]^T$. Then

$$\nabla f(\mathbf{x}) = \Sigma_{1\le j\le n}\ \partial/\partial x_j\ \left[\ \Sigma_{1\le i\le n}\ \ [|\mathbf{c}^{\mathrm{T}}\mathbf{x} - \mathbf{a}_i^{\mathrm{T}}\mathbf{x}|^2]\right]\mathbf{e}_j$$

$$= \Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\ \partial/\partial x_j\ [|\mathbf{c}^{\mathrm{T}}\mathbf{x} - \mathbf{a}_i^{\mathrm{T}}\mathbf{x}|^2]\right]\mathbf{e}_j$$

$$= \Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\ \partial/\partial x_j\ [|(\mathbf{c} - \mathbf{a}_i)^{\mathrm{T}}\,\mathbf{x}|^2]\right]\mathbf{e}_j$$

$$= \Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\ \partial/\partial x_j\ [((\mathbf{c} - \mathbf{a}_i)^{\mathrm{T}}\,\mathbf{x})^{\mathrm{T}}(\,(\mathbf{c} - \mathbf{a}_i)^{\mathrm{T}}\,\mathbf{x})]\right]\mathbf{e}_j$$

$$= \Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\ \partial/\partial x_j\ [\mathbf{x}^{\mathrm{T}}(\mathbf{c} - \mathbf{a}_i)(\mathbf{c} - \mathbf{a}_i)^{\mathrm{T}}\,\mathbf{x})]\right]\mathbf{e}_j$$

$$= \Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\ \partial/\partial x_j\ \left[\mathbf{x}^{\mathrm{T}}\ \left((\mathbf{c} - \mathbf{a}_i)(\mathbf{c} - \mathbf{a}_i)^{\mathrm{T}}\right)\mathbf{x}\right]\mathbf{e}_j.\right.$$

Using the results from assignment 2, as follows:

Letting $\mathbf{c} - \mathbf{a}_i = \mathbf{b}_i$, then

$$\Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\ \partial/\partial x_j\ \left[\mathbf{x}^{\mathrm{T}}\ \left((\mathbf{c} - \mathbf{a}_i)(\mathbf{c} - \mathbf{a}_i)^{\mathrm{T}}\right)\mathbf{x}\right]\mathbf{e}_j\right.$$

$$= \Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\partial/\partial x_j\ \mathbf{x}^{\mathrm{T}}\left(\mathbf{b}_i\mathbf{b}_i^{\mathrm{T}}\right)\mathbf{x}\right]\mathbf{e}_j.$$

From the results in problem 1(b), assignment 2, $BB^{\mathrm{T}} = \Sigma_{1\le i\le n}\ \mathbf{b}_i\mathbf{b}_i^{\mathrm{T}}$.

So,

$$\Sigma_{1\le j\le n}\,\Sigma_{1\le i\le n}\left[\partial/\partial x_j\ \mathbf{x}^{\mathrm{T}}\left(\mathbf{b}_i\mathbf{b}_i^{\mathrm{T}}\right)\mathbf{x}\right]\mathbf{e}_j = \Sigma_{1\le j\le n}\ \partial/\partial x_j\left[\mathbf{x}^{\mathrm{T}}BB^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j.$$

The result from problem 1(c), assignment 2 also states that $\mathbf{x}^{\mathrm{T}}BB^{\mathrm{T}}\mathbf{x}$ is symmetric, and positive semi-definite. Therefore, all of the eigenvalues of BB^{T} are real.

Now, taking the gradient of g, gives

$$\nabla g(\mathbf{x}) = \Sigma_{1\le j\le n}\ \partial/\partial x_j\left[\mathbf{x}^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j,$$

where $\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1$. Employing the Lagrange multiplier, and solving the system for lambda yields

$$\nabla f(\mathbf{x}) = \lambda\nabla g(\mathbf{x})$$

$$\Rightarrow\ \Sigma_{1\le j\le n}\ \partial/\partial x_j\left[\mathbf{x}^{\mathrm{T}}BB^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j = \lambda\Sigma_{1\le j\le n}\ \partial/\partial x_j\left[\mathbf{x}^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j.$$

$$\Rightarrow\ \int\Sigma_{1\le j\le n}\ \partial/\partial x_j\left[\mathbf{x}^{\mathrm{T}}BB^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j\ dx_j = \int\lambda\Sigma_{1\le j\le n}\ \partial/\partial x_j\left[\mathbf{x}^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j\ dx_j$$

$$\Rightarrow \ \Sigma_{1\le j\le n} \int \partial/\partial x_j \left[\mathbf{x}^{\mathrm{T}}BB^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j \, dx_j = \lambda\Sigma_{1\le j\le n} \int \partial/\partial x_j \left[\mathbf{x}^{\mathrm{T}}\mathbf{x}\right]\mathbf{e}_j \, dx_j$$

$$\Rightarrow \ \mathbf{x}^{\mathrm{T}}BB^{\mathrm{T}}\mathbf{x} = \lambda\mathbf{x}^{\mathrm{T}}\mathbf{x},$$

Left multiplying by **x,** and noting that $\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1 \Rightarrow \mathbf{x}\mathbf{x}^{\mathrm{T}} = 1$, yields

$$(\mathbf{x}\mathbf{x}^{\mathrm{T})}BB^{\mathrm{T}}\mathbf{x} = \lambda(\mathbf{x}\mathbf{x}^{\mathrm{T}})\mathbf{x} \ \ \Rightarrow \ BB^{\mathrm{T}}\mathbf{x} = \lambda\mathbf{x},$$

where **x** is the eigenvector to some real eigenvalue, lambda. To find lambda, left multiply the last equation by $\mathbf{x}^{\mathrm{T}}$ to get

$$\mathbf{x}^{\mathbf{T}}BB^{\mathrm{T}}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\lambda\mathbf{x} = \ \lambda(\mathbf{x}^{\mathrm{T}}\mathbf{x}) = \lambda.$$

So, to minimize

$$\min_{\mathbf{x},\mathbf{y}} \left\{\Sigma_{1\le i\le n} \ |\mathbf{a}_i - \mathbf{y}|^2 - |\mathbf{x}^{\mathrm{T}}\mathbf{a}_i|^2, \text{ subject to } \mathbf{y}^{\mathrm{T}}\mathbf{x} = 0,\ \mathbf{x}^{\mathrm{T}}\mathbf{x} = 1\right\},$$

x is an eigenvector, with the corresponding real eigenvalue $\lambda = \mathbf{x}^{\mathrm{T}}BB^{\mathrm{T}}\mathbf{x}$, of the matrix BB^{T}.