Bruce James

Assignment 4:

1.
(a) Using the method of Lagrange multipliers[i] to find $\mathbf{y}$, such that

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

let $f(\mathbf{y})$ be the vector-valued objective function $f(y_1,\dots,y_n) = \sum_{1\le i\le n} |\mathbf{a}_i - \mathbf{y}|^2$, and $g(\mathbf{y}) = \mathbf{y}^T\mathbf{x} = 0$ be the constraint function. Taking the gradient of f yields

$$\nabla f(\mathbf{y}) = \textstyle\sum_{1\le j\le n} f_j(y_1,\dots,y_n)\, \mathbf{e}_j,$$

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 $\mathbf{e}_j=[0,\dots,\mathbf{e}_j=1,\dots,0]^T$. With $|\mathbf{a}_i - \mathbf{y}|^2 = \mathbf{a}_i^T\mathbf{a}_i - 2\mathbf{a}_i^T\mathbf{y} + \mathbf{y}^T\mathbf{y}$ for each i, then

$$\nabla f(\mathbf{y}) = \textstyle\sum_{1\le j\le n} \partial/\partial y_j \left[\sum_{1\le i\le n} |\mathbf{a}_i - \mathbf{y}|^2\right]\mathbf{e}_j$$

$$= \textstyle\sum_{1\le j\le n} \partial/\partial y_j \left[\sum_{1\le i\le n} \mathbf{a}_i^T\mathbf{a}_i - 2\mathbf{a}_i^T\mathbf{y} + \mathbf{y}^T\mathbf{y}\right]\mathbf{e}_j$$

Since we are in $\mathbf{R}^n$, and the 2-norm is an assignment from $\mathbf{R}^n$ to $\mathbf{R}$ ($|\mathbf{a}_i - \mathbf{y}|^2$ is the Euclidean norm assigning the n-vector $\mathbf{y}$ to a real number), then from a theorem in real analysis[ii], f is continuous. So, from another property of analysis, we can interchange the summand with partial derivative, yielding

$$\nabla f(\mathbf{y}) = \textstyle\sum_{1\le j\le n} \sum_{1\le i\le n} \partial/\partial y_j \left[\mathbf{a}_i^T\mathbf{a}_i - 2\mathbf{a}_i^T\mathbf{y} + \mathbf{y}^T\mathbf{y}\right]\mathbf{e}_j$$

$$= \textstyle\sum_{1\le j\le n} \sum_{1\le i\le n} \partial/\partial y_j \left[\mathbf{a}_i^T\mathbf{a}_i - 2(a_{ji}y_j) + y_j^2\right]\mathbf{e}_j$$

$$= \textstyle\sum_{1\le j\le n} \left[\sum_{1\le i\le n} (-2a_{ji} + 2y_j)\, \mathbf{e}_j\right]$$

$$= \textstyle\sum_{1\le j\le n} \left[\sum_{1\le i\le n} (-a_{ji})\, \mathbf{e}_j + \sum_{1\le i\le n} y_j\, \mathbf{e}_j\right]$$

$$= 2\left[\textstyle\sum_{1\le i\le n} (-\mathbf{a}_i) + \sum_{1\le i\le n} \mathbf{y}\right]$$

$$= 2n\mathbf{y} - 2\textstyle\sum_{1\le i\le n} \mathbf{a}_i.$$

Now, taking the gradient of g, gives

$$\nabla g(\mathbf{y}) = \textstyle\sum_{1\le j\le n} \partial/\partial y_j\, [y_j x_j]\, \mathbf{e}_j$$

$$\nabla g(\mathbf{y}) = \partial/\partial y_j\, [y_1x_1 + \dots + y_nx_n]\mathbf{e}_j$$

$$= \Sigma_{1\le j\le n}\ (x_j)\mathbf{e}_j$$

$$= \mathbf{x}.$$

Employing the Lagrange multiplier, and solving the system for lambda yields

$$\nabla f(\mathbf{y}) = \lambda\nabla g(\mathbf{y}) \quad \Rightarrow \quad 2n\mathbf{y} - 2\,\Sigma_{1\le i\le n}\,\mathbf{a}_i = \lambda\,\mathbf{x}$$

Now, left-multiplying by $\mathbf{x}^T$, and using the conditions $\mathbf{y}^T\mathbf{x} = 0 \Rightarrow \mathbf{x}^T\mathbf{y} = 0$, and $\mathbf{x}^T\mathbf{x} = 1$,

$$\mathbf{x}^T\,(2n\mathbf{y} - 2\,\Sigma_{1\le i\le n}\,\mathbf{a}_i) = \mathbf{x}^T\,\lambda\,\mathbf{x}$$

$$\Rightarrow \quad \mathbf{x}^T\,(2n)\mathbf{y} - \mathbf{x}^T\,(2)\,\Sigma_{1\le i\le n}\,\mathbf{a}_i = \lambda\,(\mathbf{x}^T\,\mathbf{x})$$

$$\Rightarrow \quad (2n)\mathbf{x}^T\mathbf{y} - 2\mathbf{x}^T\,\Sigma_{1\le i\le n}\,\mathbf{a}_i = \lambda\,(\mathbf{x}^T\,\mathbf{x})$$

$$\Rightarrow \quad \lambda = -2\mathbf{x}^T\,\Sigma_{1\le i\le n}\,\mathbf{a}_i.$$

Substituting for lambda gives

$$2n\mathbf{y} - 2\,\Sigma_{1\le i\le n}\,\mathbf{a}_i = -2\mathbf{x}^T\,(\Sigma_{1\le i\le n}\,\mathbf{a}_i)\mathbf{x}$$

$$\Rightarrow \quad \mathbf{y} = 1/n\,[\Sigma_{1\le i\le n}\,\mathbf{a}_i - \mathbf{x}^T\,(\Sigma_{1\le i\le n}\,\mathbf{a}_i)\mathbf{x}],$$

with $\mathbf{x}$ and its transpose as known entities. So, to minimize (2),

$$\mathbf{y} = 1/n\,[\Sigma_{1\le i\le n}\,\mathbf{a}_i - \mathbf{x}^T\,(\Sigma_{1\le i\le n}\,\mathbf{a}_i)\mathbf{x}].$$

---

[i] Larson, *Calculus: Early Transcendental Functions*, 6th Ed.
[ii] Marsden, Hoffman, *Elementary Classical Analysis*, 2nd Ed.