PS #4 — Least squares problems

For this problem set, you must complete Problems 1-5 as well as one of either Problems 6 or 7. Problem 8 is a bonus problem.

Problem 1 (Datta 7.1). Let $\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$ be a Householder reflector (recall that $\mathbf{u}$ has unit norm). Show that $\mathbf{H}\mathbf{u} = -\mathbf{u}$, and that $\mathbf{H}\mathbf{v} = \mathbf{v}$ if $\mathbf{v}^T\mathbf{u} = 0$. Interpret these two facts geometrically.

Problem 2 (Datta 7.2). Let $\mathbf{x} \in \mathbb{R}^n$. Describe a method for computing a Householder reflector $\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$ such that $\mathbf{H}\mathbf{x}$ has zeros in the entries $r+1$ through n, i.e., $x_{r+1} = x_{r+2} = \cdots = x_n = 0$. Given $\mathbf{x} = [1,2,3]^T$, apply your method to find $\mathbf{H}$ such that $\mathbf{H}\mathbf{x}$ has a zero in its last entry.

Problem 3 (Datta 7.8). Given $\mathbf{v} = [1,2,3]^T$, find a Givens rotation matrix $G(c,s,1,3)$ that zeros out the third component of $\mathbf{v}$, i.e., the third component of $G(c,s,1,3)\mathbf{v}$ is zero.

Problem 4. Let $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ denote the economic QR factorization of $\mathbf{A} \in \mathbb{R}^{m\times n}$ with $\text{rank}(\mathbf{A}) = n$. *Part (a):* Show that $\hat{\mathbf{R}}^{-1}\hat{\mathbf{Q}}^T = \mathbf{A}^\dagger$ by showing that $\hat{\mathbf{R}}^{-1}\hat{\mathbf{Q}}^T$ satisfies the four parts of the definition of the pseudoinverse. *Part (b): (AG 6.4.5 (b))* The least squares solution to

$$\underset{\mathbf{x}\in\mathbb{R}^n}{\arg\min}\ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2, \tag{1}$$

is given by $\mathbf{x} = \mathbf{A}^\dagger\mathbf{b} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} = \cdots = \hat{\mathbf{R}}^{-1}\hat{\mathbf{Q}}^T\mathbf{b}$. The expression $\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$ involves a linear system solve with condition number $[\kappa(\mathbf{A})]^2$, while the expression $\mathbf{x} = \hat{\mathbf{R}}^{-1}\hat{\mathbf{Q}}^T\mathbf{b}$ involves a linear system with condition number $\kappa(\mathbf{A})$. Where is the improvement step hidden? Explain.

Problem 5. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ with $\text{rank}(\mathbf{A}) = n$ and let $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ be the economic QR factorization. Show that

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 = \|\hat{\mathbf{R}}\mathbf{x} - \hat{\mathbf{Q}}^T\mathbf{b}\|_2^2 + \|(\mathbf{I} - \hat{\mathbf{Q}}\hat{\mathbf{Q}}^T)\mathbf{b}\|_2^2 \tag{2}$$

and that

$$\underset{\mathbf{x}\in\mathbb{R}^n}{\arg\min}\ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 = \underset{\mathbf{x}\in\mathbb{R}^n}{\arg\min}\ \|\hat{\mathbf{R}}\mathbf{x} - \hat{\mathbf{Q}}^T\mathbf{b}\|_2^2. \tag{3}$$

Problem 6*. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{L} \in \mathbb{R}^{k\times n}$ be rectangular matrices satisfying $\ker(\mathbf{A}) \cap \ker(\mathbf{L}) = \{\mathbf{0}_n\}$. Additionally, let $\mathbf{K}$ be a matrix with orthonormal columns such that $\ker(\mathbf{L}) = \mathrm{col}(\mathbf{K})$. Using the four parts of the definition of the pseudoinverse, prove the identity

$$\left(\mathbf{A}(\mathbf{I} - \mathbf{L}^\dagger\mathbf{L})\right)^\dagger = \mathbf{K}(\mathbf{A}\mathbf{K})^\dagger. \tag{4}$$

Make sure to justify all steps. Why is the assumption that $\mathbf{K}$ has orthonormal columns important here? This identity has applications in standard form transformations for regularized (penalized) least squares problems arising in inverse problems, where it can be used to simplify computations with the oblique ($\mathbf{A}$-weighted) pseudoinverse

$$\mathbf{L}_{\mathbf{A}}^\dagger = (\mathbf{I} - (\mathbf{A}(\mathbf{I} - \mathbf{L}^\dagger\mathbf{L}))^\dagger\mathbf{A})\mathbf{L}^\dagger. \tag{5}$$

Hints: some facts that you may find helpful to use (with justification) are that (i) $\mathbf{A}\mathbf{K}$ *has linearly independent columns, (ii)* $\mathbf{K}\mathbf{K}^T = \mathbf{I} - \mathbf{L}^\dagger\mathbf{L}$, and (iii) $\mathbf{L}^\dagger\mathbf{L} = \mathbf{I} - \mathbf{K}\mathbf{K}^T$.

Problem 7*. For a symmetric matrix $\mathbf{C} \in \mathbb{R}^{p\times p}$ and a tall matrix $\mathbf{X} \in \mathbb{R}^{p\times q}$ with $p \geq q$, a generalization of Ostrowski's theorem states that the eigenvalues of $\mathbf{X}^T\mathbf{C}\mathbf{X}$ satisfy

$$\lambda_{i+(p-q)}(\mathbf{C})\lambda_q(\mathbf{X}^T\mathbf{X}) \leq \lambda_i(\mathbf{X}^T\mathbf{C}\mathbf{X}) \leq \lambda_i(\mathbf{C})\lambda_1(\mathbf{X}^T\mathbf{X}), \quad i = 1, \ldots, q. \tag{6}$$

Use this result to explain why the condition number of the projected least squares problem

$$\underset{\mathbf{x}\in\mathrm{col}(\mathbf{V})}{\arg\min}\ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 = \mathbf{V}\left(\underset{\mathbf{z}\in\mathbb{R}^d}{\arg\min}\ \|\mathbf{A}\mathbf{V}\mathbf{z} - \mathbf{b}\|_2\right) \tag{7}$$

can be no worse than that of the unconstrained least squares problem

$$\underset{\mathbf{x}\in\mathbb{R}^n}{\arg\min}\ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2, \tag{8}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{V} \in \mathbb{R}^{n\times d}$ has orthonormal columns with $d \leq n$. In other words, projecting the least squares problem onto an orthonormal basis for any subspace of $\mathbb{R}^n$ cannot worsen the conditioning — in general, the conditioning may worsen if the basis is not an orthonormal basis.

Problem 8 (bonus). Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and suppose that $\mathrm{rank}(\mathbf{A}) = r < n$. The (pivoted) QR factorization of the rank-deficient matrix $\mathbf{A}$ is given by

$$\mathbf{A}\mathbf{P} = \mathbf{Q}\begin{bmatrix}\mathbf{R}_{11} & \mathbf{R}_{12}\\ 0 & 0\end{bmatrix} \tag{9}$$

where $\mathbf{R}_{11} \in \mathbb{R}^{r\times r}$ is upper triangular and nonsingular, $\mathbf{P}$ is a permutation matrix, and $\mathbf{Q} \in \mathbb{R}^{m\times m}$ is an orthogonal matrix. Using Householder reflectors from the right, this factorization can be converted to one of the form

$$\mathbf{A}\mathbf{P} = \mathbf{Q}\begin{bmatrix}\tilde{\mathbf{R}} & 0\\ 0 & 0\end{bmatrix}\mathbf{V}^T \tag{10}$$

where $\tilde{\mathbf{R}} \in \mathbb{R}^{r\times r}$ is nonsingular and $\mathbf{V} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix, which is known as the *complete orthogonal decomposition*. Show that

$$\mathbf{A}^\dagger = \mathbf{P}\mathbf{V}\begin{bmatrix}\tilde{\mathbf{R}}^{-1} & 0\\ 0 & 0\end{bmatrix}\mathbf{Q}^T \tag{11}$$

by verifying the four conditions in the definition of the pseudoinverse.