Homework 6

Due date: April 27, 2020, 11:55PM.

Bibliography: Trefethen & Bau, Lectures 36-37

The first 4 questions are your normal homework. The final three questions are your final examination. Honor system applies. As usual, you are free to discuss among yourselves approaches to homework questions. The final examination questions are open-book, but you should draft your answer without discussing approaches with anyone else.

1 Homework questions

1. Ex. 36.2, p. 283.
2. Ex. 36.3, p. 284.
3. Ex. 36.4, p. 284 (modify Lesson20.ipynb to plot eigenvalues and Ritz estimates, and lemniscates also).
4. Ex. 38.6, p. 302 (the code framework for this problem will be developed in class).

2 Final examination questions

Note: The following questions do not require extensive calculation, but rather careful, insightful application of concepts from linear algebra and scientific computation. You should be able to draft answers to these questions in the date interval 04/23 - 04/27.

1. Ex. 39.5, p.312
2. Ex. 40.2, p.319
3. Consider m artificial neurons that each linearly combine n common inputs $\boldsymbol{x} \in [0,1]^n$ into an output vector $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^m$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, subsequently thresholded into a binary synapse vector $\boldsymbol{b} = \boldsymbol{\varphi}(\boldsymbol{y}) \in \mathbb{R}^m$, with $\boldsymbol{\varphi}$ the componentwise Heaviside function, i.e., $b_i = \varphi(y_i)$, $i = 1, \ldots, m$, $\varphi(s) = 1$ if $s > 0$, $\varphi(s) = 0$ if $s \leqslant 0$. The overall transformation $\boldsymbol{b} = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\varphi}(\boldsymbol{A}\boldsymbol{x})$ is nonlinear, but insight into the behavior of the layer of artificial neurons can still be obtained through concepts from linear algebra. Let $\mathcal{R}, \mathcal{N}$ denote the range, null space operators, i.e., $\mathcal{R}(\boldsymbol{f}(\mathcal{S}))$ is the set of values attained by $\boldsymbol{f}$ for all function argument values within $\mathcal{S}$, and $\mathcal{N}(\boldsymbol{f}(\mathcal{S}))$ is the set of elements of $\boldsymbol{s} \in \mathcal{S}$ for which $\boldsymbol{f}(\boldsymbol{s}) = \boldsymbol{0}$. If $\mathcal{S}$ is not specified the function entire domain is assumed, i.e., $\boldsymbol{f}: \mathcal{D} \to \mathcal{C}$, $\mathcal{R}(\boldsymbol{f}) = \mathcal{R}(\boldsymbol{f}(\mathcal{D})) \subseteq \mathcal{C}$, $\mathcal{N}(\boldsymbol{f}) = \mathcal{N}(\boldsymbol{f}(\mathcal{D})) \subseteq \mathcal{D}$.

   Recall that the singular value decomposition of $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ provides bases for the four fundamental spaces $\mathcal{R}(\boldsymbol{A})$, $\mathcal{N}(\boldsymbol{A})$, $\mathcal{R}(\boldsymbol{A}^T)$, $\mathcal{N}(\boldsymbol{A}^T)$, and the rank-nullity theorem specifies

   $$\dim \mathcal{R}(\boldsymbol{A}) + \dim \mathcal{N}(\boldsymbol{A}^T) = m, \dim \mathcal{R}(\boldsymbol{A}^T) + \dim \mathcal{N}(\boldsymbol{A}) = n, \operatorname{rank}(\boldsymbol{A}) = r = \dim \mathcal{R}(\boldsymbol{A}) = \dim \mathcal{R}(\boldsymbol{A}^T).$$

   a) Characterize $\mathcal{R}(\boldsymbol{f}), \mathcal{N}(\boldsymbol{f})$ in terms of $\mathcal{R}(\boldsymbol{A})$, $\mathcal{N}(\boldsymbol{A})$, $\mathcal{R}(\boldsymbol{A}^T)$, $\mathcal{N}(\boldsymbol{A}^T)$.

   b) Invoke concepts from solving the least squares problem $\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|$ to solve the problem

   $$\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x})\|.$$

   c) Assume that $\boldsymbol{A}$ is square, $\boldsymbol{A} \in \mathbb{R}^{m \times m}$, and orthogonally diagonalizable, $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T$, $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$, with N eigenvalues λ_i, having algebraic, geometric multiplicities (m_i, n_i), $i = 1, \ldots, N$, and associated eigenspaces $\mathcal{E}_i$. Characterize $\mathcal{R}(\boldsymbol{f}(\mathcal{E}_i))$, $\mathcal{N}(\boldsymbol{f}(\mathcal{E}_i))$.