Linear Algebra for Computer Science
Homework 4

Read these first:

i You may write your solutions on paper, under a word processing software (MS-word, Libre Office, etc.), or under LaTeX.

ii If writing on paper, you must use a scanner device or a Camera Scanner (CamScanner) software to scan the document and submit a *single* PDF file.

iii Up to 15% extra score will be given to solutions written under LaTeX, provided that you follow either of the following conventions:

(a) Represent scalars with normal (italic) letters (a, A), vectors with bold lower-case letters ($\mathbf{a}$, using `\mathbf{a}`), and matrices with bold upper-case letters ($\mathbf{A}$, using `\mathbf{A}`), or

(b) represent scalars with normal (italic) letters (a, A), vectors with bold letters ($\mathbf{a}$, $\mathbf{A}$), and matrices with typewriter upper-case letters ($\mathtt{A}$, using `\mathtt{A}`).

(c) You latex document must contain a *title*, a *date*, and your name as the author.

(d) In all cases, you must submit a *single* PDF file.

(e) If writing under LaTeX, you must submit the *.tex* source (and other necessary source files if there are any) in addition to the PDF file.

Here is a short tutorial on LaTeX: `https://www.overleaf.com/learn/latex/Learn_LaTeX_in_30_minutes`

Questions

Projections

1. Consider a linear subspace $\mathcal{S}$ and a vector $\mathbf{y} \in \mathcal{S}$. Using the projection formula, show that the projection of $\mathbf{y}$ into $\mathcal{S}$ is itself.

2. For a linear subspace $\mathcal{S} \subseteq \mathbb{R}^n$ its *orthogonal complement* is defined as $\mathcal{S}^\perp = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y}^T\mathbf{x} = 0 \text{ for all } \mathbf{x} \in \mathcal{S}\}$. In other words, $\mathcal{S}^\perp$ comprises all the vectors that are perpendicular to all vectors in $\mathcal{S}$. Show that the orthogonal complement of a linear subspace is a linear subspace.

3. Prove that the *null space* of a matrix is the orthogonal complement of its *row space*.

4. Let $\mathtt{P} \in \mathbb{R}^{n\times n}$ be the projection matrix into a linear subspace $\mathcal{S}$. Show that $\mathtt{I} - \mathtt{P}$ represents the projection into the orthogonal complement of $\mathcal{S}$. Hint: First show that $\mathtt{I} - \mathtt{P}$ is a projection matrix.

Determinant

5. Prove that the determinant of an orthogonal matrix is either equal to 1 or -1.

6. Show that the determinant of a projection matrix is either equal to 0 or 1. How do you explain this geometrically?

Eigenvalues and Eigenvectors

7. What is the relationship between the eigenvalues and eigenvectors of the square matrix $\mathtt{A}$ and those of $\mathtt{A} - \alpha\mathtt{I}$ where $\alpha \in \mathbb{R}$ and $\mathtt{I}$ is the identity matrix?

8. Prove that any eigenvalue of $\mathtt{A}$ is also an eigenvalue of $\mathtt{A}^T$. (Hint: use the characteristic polynomial).

9. The square matrix $\mathtt{A}$ is called (left) stochastic (or a Markov matrix) if its elements are nonnegative and its columns add up to 1 (programmatically sum(A,axis=0) == ones((1,n))). Prove that $\mathtt{A}$ has at least one unit eigenvalue $\lambda = 1$. (Hint: First prove that $\mathtt{A}^T$ has a unit eigenvalue.)

10. Let $\mathbf{v}$ be an eigenvector of $\mathtt{A}$ with a nonzero corresponding eigenvalue $\lambda \neq 0$. Prove that

    (a) $\mathbf{v}$ is in the column space of $\mathtt{A}$.

    (b) The (orthogonal) projection of $\mathbf{v}$ into the row space of $\mathtt{A}$ is nonzero. (Hint: decompose the vector as $\mathbf{v} = \mathbf{v}_r + \mathbf{v}_n$ where $\mathbf{v}_r$ and $\mathbf{v}_n$ are in the row space and null space of $\mathtt{A}$, respectively. Then show that $\mathbf{v}_r$ is nonzero)

11. Let $\mathtt{A}$ be a symmetric real matrix with real eigenvalues $1, 2, \ldots, n$, and corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$. Prove that if $\lambda_i \neq \lambda_j$ then $v_i \perp v_j$.

Positive Definite Matrices

For all question in this section, by *positive definite* we mean *symmetric positive definite*.

12. Prove that a symmetric matrix is positive definite if and only if all its eigenvalues are positive. (Remember from the class that the eigen-decomposition of a symmetric matrix is in the form of $\mathtt{A} = \mathtt{V}\Lambda\mathtt{V}^{-1} = \mathtt{V}\Lambda\mathtt{V}^T$.)

13. Show that the diagonal elements of a positive definite matrix are all positive.

14. Remember that an operation $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ defined on a vector space $\mathcal{V}$ is an *inner product* if

    (a) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ for all $\mathbf{u} \in \mathcal{V}$,

    (b) $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$,

    (c) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$,

    (d) $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$.

    Let $\mathtt{A} \in \mathbb{R}^{n \times n}$ be any *positive definite* matrix. Show that the operation $\langle \cdot, \cdot \rangle_\mathtt{A} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

    $$\langle \mathbf{u}, \mathbf{v} \rangle_\mathtt{A} = \mathbf{u}^T \mathtt{A} \mathbf{v}$$

    is indeed an inner product.

Singular Value Decomposition

15. Let $\mathtt{A}$ be a nonsingular square matrix and $A = \mathtt{U}\Sigma\mathtt{V}^T$ be its (full) SVD. Prove that $\det(\mathtt{U})\det(\mathtt{V}) = \text{sign}(\det(\mathtt{A}))$, that is $\det(\mathtt{U})\det(\mathtt{V}) = 1$ if $det(\mathtt{A}) > 0$ and $\det(\mathtt{U})\det(\mathtt{V}) = 1$ if $\det(\mathtt{A}) < 0$.

16. Show that for a symmetric positive definite matrix the eigenvalue decomposition $\mathtt{A} = \mathtt{V}\Lambda\mathtt{V}^{-1} = \mathtt{V}\Lambda\mathtt{V}^T$ is the same as its singular value decomposition.

17. Find a way to obtain the SVD of a symmetric matrix from its eigenvalue decomposition $\mathtt{A} = \mathtt{V}\Lambda\mathtt{V}^T$. Notice that the diagonal elements of Λ might be negative.

18. Consider a matrix $\mathtt{A} \in \mathbb{R}^{m \times n}$ and two orthogonal matrices $\mathtt{P} \in \mathbb{R}^{m \times m}$ and $\mathtt{Q} \in \mathbb{R}^{n \times n}$. Show that the singular values of $\mathtt{PAQ}$ is the same as the singular values of $\mathtt{A}$.

Multivariate Calculus

19. Show that for a matrix $\mathtt{A} \in \mathbb{R}^{n\times n}$ the gradient of the expression $\mathbf{x}^T \mathtt{A}\mathbf{x}$ is equal to $(\mathtt{A} + \mathtt{A}^T)\,\mathbf{x}$. What is the gradient when $\mathtt{A}$ is symmetric?

20. Show that for a symmetric matrix $\mathtt{B}$ the gradient of $1/(\mathbf{x}^T \mathtt{B}\mathbf{x})$ with respect to $\mathbf{x}$ is $-2\mathtt{B}\mathbf{x}/(\mathbf{x}^T \mathtt{B}\mathbf{x})^2$ (if the gradient exists at $\mathbf{x}$).

21. Show that for symmetric matrices $\mathtt{A}$ and $\mathtt{B}$ the gradient of $f(\mathbf{x}) = (\mathbf{x}^T \mathtt{A}\mathbf{x})/(\mathbf{x}^T \mathtt{B}\mathbf{x})$ with respect to x is equal to

$$2(\mathtt{A}\mathbf{x}(\mathbf{x}^T \mathtt{B}\mathbf{x}) - \mathtt{B}\mathbf{x}(\mathbf{x}^T \mathtt{A}\mathbf{x}))/(\mathbf{x}^T \mathtt{B}\mathbf{x})2 = 2(\mathtt{A}\mathbf{x} - f(\mathbf{x})\,\mathtt{B}\mathbf{x})/(\mathbf{x}^T \mathtt{B}\mathbf{x}),$$

if the gradient exists at $\mathbf{x}$.

22. Let $\mathtt{A}$ be symmetric. Calculate the gradient of $\exp(-\mathbf{x}^T \mathtt{A}\mathbf{x})$ with respect to $\mathbf{x}$.

23. Let $\mathtt{A}$ be (symmetric) positive definite. Compute the gradient of $\log(1 + \mathbf{x}^T \mathtt{A}\mathbf{x})$ with respect to $\mathbf{x}$.

24. Consider the function $f(\mathbf{x}) = \mathbf{x}^T \mathtt{A}\mathbf{x}/\left\|\mathbf{x}\right\|^2 = \mathbf{x}^T \mathtt{A}\mathbf{x}/(\mathbf{x}^T\mathbf{x})$ defined for a symmetric matrix $\mathtt{A}$. Show that the critical points of f are exactly the eigenvectors of $\mathtt{A}$. The critical points of a function f are points $\mathbf{x}$ at which the gradient is zero or nonexistant.

25. Consider the function $f(\mathbf{x}) = \mathbf{x}^T \mathtt{A}\mathbf{x}/(\mathbf{x}^T \mathtt{B}\mathbf{x})$ defined for symmetric matrices $\mathtt{A}$ and $\mathtt{B}$. Show that if $\mathtt{B}$ is invertible then the critical points of f are either the points for which $\mathbf{x}^T \mathtt{B}\mathbf{x} = 0$ or the eigenvectors of $\mathtt{B}^{-1}\mathtt{A}$.