MATEMATISKA INSTITUTIONEN
STOCKHOLMS UNIVERSITET
Avd. Matematik
Examinator: Yishao Zhou

Tentamensskrivning i
Linear Algebra and
Learning from Data AN, 7,5 hp
October 25, 2022

You are allowed to bring an A4 page (double sides) with whatever you think is important. A possible dictionary German-English, English-German, e.g. The Oxford-Duden German dictionary is allowed.
You must justify/motivate well your arguments.
Follow carefully the instructions. You will not be granted extra points if you choose to do more than required in this problem set.

---

1. (i) Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite. Show that $B^tAB > 0$ if and only if the null space $\mathcal{N}(B) = \{0\}$, where $B \in \mathbb{R}^{n\times k}$.

   (ii) Let $\langle A, B\rangle = \mathrm{tr}(A^tB)$ for any $A, B \in \mathbb{R}^{n\times n}$. Show that this defines an inner product on the vector space $\mathbb{R}^{n\times n}$, and the Frobenius norm is the induced norm of this inner product.

   12 p

2. Let $A \in \mathbb{R}^{n\times n}$ and we define the induced matrix norm: $\|A\| = \max_{x\neq 0} \frac{\||Ax\|}{\|x\|}$ where $\|\cdot\|$ is any vector norm on $\mathbb{R}^n$.

   (i) Justify the well-definedness of this definition.

   (ii) Show that $\|A\|_2$, using the vector norm $\|\cdot\|_2$, is the largest singular value of A.

   (iii) Let $A = I$ be the $n \times n$ identity matrix. Determine $\|I\|_2$, $\|I\|_F$ and $\|I\|_N$. Next let $A = Q$ be an orthogonal matrix. Determine $\|Q\|_2$, $\|Q\|_F$ and $\|Q\|_N$. As a reminder F and N stand for Frobenius and nuclear, respectively.

   (iv) Find the relations between these three norms for any square matrix A. Show also that $\|A\|_F = \|A\|_2$ if A is a rank 1 matrix.

   (v) Define $\kappa = \|A\|\|A^{-1}\|$, the condition number that measures conditioning of the matrix A in solving $Ax = b$. However, in the least square problems the matrix A is $m\times n$ and $m \geq n$. Modify the definition of $\kappa(A)$ using $\|\cdot\|_2$ which will be the same as defined before if $m = n$.

   12 p

3. Given the diagonal matrix $A = \mathrm{diag}(4, 3, 2, 1)$. What is the best 2-rank approximation of A (in which sense)? State the general result for the approximation of any A with low rank. Give some application areas of this theorem.

   12 p

4. (i) Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ given by $f(x) = x^tAx$ where $A = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 2 & \theta \end{pmatrix}$. Find H, the Hessian of f without computing the partial derivatives. For what values of θ is f strictly convex?

   (ii) Argue that the matrix H can be written as a sum of rank one matrices.

   (iii) What is the smallest eigenvalue of H, without computing, if $\theta = 2$?

   12 p

5. Let $A = \begin{pmatrix} B & b \\ b^t & a \end{pmatrix} \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{(n-1)\times(n-1)}$ be symmetric and $b \in \mathbb{R}^{n-1}$. Assume that A has eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ and B has eigenvalues $\mu_1 \leq \cdots \leq \mu_{n-1}$. Show that

$$\lambda_1 \leq \mu_1 \leq \lambda_2 \leq \mu_2 \leq \cdots \leq \mu_{n-1} \leq \lambda_n.$$

   12 p

You have finished the exam if the point of your homework is $p_h \geq 24$ and your project part is graded as Pass. Continue otherwise. Now if you don't collected enough points, follow carefully the instructions below.

6. (*If you passed the projects, go to next.*)
   Fisher's LDA attempts to find a separation vector onto which the projection of different classes are "best separated" by solving the optimization problem $\max_{\|v\|\neq 0} \frac{(v^tm_A - v^tm_B)^2}{v^t(\Sigma_A+\Sigma_B)v}$ where m_C, Σ_C are sampled mean and covariance matrices for $C \in \{A, B\}$ the two classes. Find an optimal solution. Argue how you will deal with the situation where $\Sigma_A + \Sigma_B$ is not positive definite or this matrix is nearly singular.

   10 p

You have finished the exam if the point of your homework is $p_h \geq 24$. Continue otherwise.

7. (i) Given two n-vectors a and x, define their circular convolution $y = a * x$ as $y_y = \sum_{l=0}^{n-1} a_{k-l} x_l$, where the indices in the sum are evaluated modulo n. Show that the circular convolution is commutative and associative.

   (ii) Assume that the matrix A has simple eigenvalues. Show that A and B are simultaneously diagonalizable if and only if they commute. In this case the diagonalizing basis is made up of the eigenvectors of A.

   (iii) Let S and its adjoint S^* be the circular shift operators defined by $S(x_0, , , , , x_{n-1}, x_n) = (x_{n-1}, x_0, ..., x_{n-2})$ and $S^*(x_0, ..., x_{n-1}, x_n) = (x_1, , ..., x_{n-1}, x_0)$, respectively. Show that any matrix M that commutes with the circular shift operator S must be a circulant matrix.

   (iv) Find all eigenvalues of S^* and their corresponding eigenvectors. Justify that the operator S^* on $\mathbb{R}^n$ has n distinct eigenvalues.

   (v) Show that any circulant matrix C has the same eigenvectors as those of S^*.

   10 p

You have finished the exam if your homework $23 \geq p_h \geq 16$. Continue otherwise.

8. Let $A \in \mathbb{R}^{m \times n}$ with $m > n$. Consider the equation $Ax = b$.

   (i) Show how you derive a solution if $A^t A$ is not invertible.

   (ii) Describe the gradient descent method for solving the least square problem $\min \|b - Ax\|_2^2$ assuming A has full column rank.

   (iii) Find the conditions for the convergence of this method and derive the convergence rate.

   10 p

You have finished the exam if your homework $16 \geq p_h \geq 8$. Continue otherwise.

9. How do you solve the real polynomial equation $p(s) = s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0$ using linear algebra? 10 p

---

To get the graded paper fill in the formula at `https://survey.su.se/Survey/44514/sv` *(Svenska) or* `https://survey.su.se/Survey/44514/en` *(English)*