- The exam has 2 pages with 9 problems.
- Please justify your answers, proving the statements you make. You are allowed to refer to results shown in lectures/recitations/homeworks as long as you state them precisely, meaning that you should say exactly which hypothesis are needed in the result you use.
- Partial answers will be graded. But not justified answers will not necessarily yield credit.
- Put your name and your netID on your paper. Number the pages on both sides with $1/x, 2/x$ etc... with x your total final number of pages.
- Once I announce that time is up, drop your pen immediately. Then you will be allowed to take out your phone to take a picture of all the pages of your exam before handing it to me. You will have to upload your pictures to Gradescope before the end of the day.

Problem 0.1 (6 points)**.** *Correct/Incorrect.* ***Justify***, *if Incorrect a counter-example suffices. Consider $A \in \mathbb{R}^{n\times m}$ with $n > m$ and $B \in \mathbb{R}^{k\times k}$.*

(a) *If the columns of A form an orthonormal family, $A^{-1} = A^\top$.*

(b) *Suppose v_1 and v_2 are right singular vectors of A associated with the same singular value, then any non-zero vector in $\mathrm{Span}(v_1, v_2)$ is also a right singular vector of A.*

(c) *There are unique matrices $U \in \mathbb{R}^{n\times n}$ orthogonal, $V \in \mathbb{R}^{m\times m}$ orthogonal and $\Sigma \in \mathbb{R}^{n\times m}$ diagonal such that $A = U\Sigma V^\top$.*

Problem 0.2 (6 points)**.** *Correct/Incorrect.* ***Justify***, *if Incorrect a counter-example suffices. Consider f and g two convex functions from $\mathbb{R}^n \to \mathbb{R}$ and*

(a) *The function $h(x_1, x_2) = x_1^2 x_2$ defined on $\mathbb{R}^2$ is convex.*

(b) *The function $h(x) = f(x)g(x)$ is always convex.*

(c) *If f is a linear transformation, the set $S = \{x \in \mathbb{R}^n | f(x) = 0\}$ is a convex set.*

Problem 0.3 (8 points)**.** *Consider $f(x, y) = (x^2 - 1)^2 + (y^2 - 1)^2$.*

(a) *Is f convex on $\mathbb{R}^2$? Justify.*

(b) *Find all the critical points of f. For each one, specify if it is a local minimum, a local maximum or a saddle point.*

(c) *Give all the solutions to the constrained optimization problem*

$$\text{minimize } f(x, y) \quad \text{subject to} \quad x \geq 1/2.$$

Is the inequality constraint active? (Hint: No need for Lagrange multipliers)

(d) *Knowing that there exits at least one, give all the solutions to the constrained optimization problem*

$$\text{minimize } f(x, y) \quad \text{subject to} \quad x + y \leq -3.$$

Is the inequality constraint active?
(Hint: You can directly use the fact that the equation $((a + 3)^2 - 1)(a + 3) = -(a^2 - 1)a$ has a unique solution for a.)

Problem 0.4 (6 points). *You are given n feature vectors $a_i \in \mathbb{R}^d$, for only a subset $m < n$ of them you also have the associated scalar y_i. You would like to find a found linear relationship between the a_i and y_i. You only have much fewer observations than features ($m < d$).*

(**a**) *First you ask yourself which are the most relevant features. Name a method you would use to select a subset d' of the d features. Which part of the data would you have used?*

(**b**) *Name an other method you could use to build a dataset with features of dimension $d' < d$ which are not necessarily a subset of the d original features. Which part of the data would you have used?*

(**c**) *You also try ridge regression of which you know a solution. For $A \in R^{m\times d}$ and $y \in R^m$,*

$$\text{minimize} \quad \frac{1}{2}\|Ax - y\|^2 + \frac{\lambda}{2}\|x\|^2 \quad \text{with solution} \quad x^{\text{ridge}} = (A^\top A + \lambda I_d)^{-1}A^\top y.$$

Using the SVD of $A = U\Sigma V^\top$, show that

$$(A^\top A + \lambda I_d)^{-1}A^\top = A^\top (AA^\top + \lambda I_m)^{-1}.$$

(Hint: Note that $\Sigma^\top\Sigma$ and $\Sigma\Sigma^\top$ are diagonal matrices.)

(**d**) *Knowing that inverting a matrix is a costly computation, considering the previous question, which expression for x^{ridge} would you choose?*

Problem 0.5 (3 points). *Given a data set of N striclty positive scalar values $x_1, \cdots x_N$ in $\mathbb{R}_*^+$, we can define the arithmetic mean M and the geometric mean m as follows*

$$M = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad m = \left(\prod_{i=1}^{N} x_i\right)^{\frac{1}{N}}.$$

Show that $M \geq m$.
(Hint: show that the exponential is convex over $\mathbb{R}$ and that you can define $y_i = log(x_i)$.)

Problem 0.6 (3 points). *Show that the map $A \to \|A\|_{\text{Sp}} = \max_{\|x\|_2=1} \|Ax\|_2 = \sigma_{\max}(A)$ (largest singular value of A) is a norm on $\mathbb{R}^{n\times m}$. In other words, show that the definition we gave in class of the spectral norm is legal.*

Problem 0.7 (3 points). *Let A a matrix in $\mathbb{R}^{n\times m}$ and $A^\dagger$ its Moore-Penrose inverse. Show that $AA^\dagger$ is the matrix of the orthogonal projection onto $\text{Im}(A)$.*

Problem 0.8 (extra - 3 points). *Find all the matrices $A \in \mathbb{R}^{n\times m}$ solutions of $\text{Tr}(AA^T) = 0$. Justify that you indeed found all the solutions.*

Problem 0.9 (extra - 5 points). *Let G be a graph with n nodes, and denote by $d_{\max}$ the maximum degree across its nodes. In other nodes, at most a node is connected to d_max other node in G. Denote by $A \in \mathbb{R}^{n\times n}$ the adjacency matrix of G.*

(**a**) *Show that if λ is an eigenvalue of A then $|\lambda| \leq d_{\max}$.*

(**b**) *Show the maximum eigenvalue of A is larger than the average degree $\bar{d} = \frac{1}{n}\sum_{k=1}^{n} \deg(k)$.*

(**c**) *Now assume that is G is a d-regular graph (d an integer), which means that all of its nodes are connected to exactly d other nodes. What is the maximum value of A? Give an associated eigenvector.*