# Kernel Methods - Practical Session № 1

AMMI 2020 18/05/2020

## Linear Algebra Recap

### Exercise 1

Let $A \in \mathbb{R}^{m \times n}$ be a real matrix. For each function $f$, **specify its codomain and compute** $\nabla f$

(a) f: $\begin{cases} \mathbb{R} \to ? \\ x \to Tr(Ax) \end{cases}$

(b) f: $\begin{cases} \mathbb{R}^n \to ? \\ \mathrm{x} \to A\mathrm{x} \end{cases}$

(c) f: $\begin{cases} \mathbb{R}^{m \times n} \to ? \\ X \to Tr(A^T X) \end{cases}$

### Exercise 2

***Reminder****: Singular Value Decomposition*

*Let $A \in \mathbb{R}^{m \times n}$. There exists two orthogonal matrices $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\Sigma \in \mathbb{R}^{m \times n}$ such that $A = U^T \Sigma V$*

Let $X \in \mathbb{R}^{n \times p}$, $\lambda > 0$.
**Prove that $M = X^T X + \lambda I_p$ is invertible.**

*Hint: Prove that the eigenvalues of $M$ are larger than $\lambda$*

## Linear Regression

Using notations from the slides:

- $Y = (y_1, ..., y_n)^T \in \mathbb{R}^n$ the vector of outcomes
- $X = (x_1, ..., x_n)^T \in \mathbb{R}^{n \times p}$ the matrix ($n$ rows=samples, $p$ columns=features)
- $\beta \in \mathbb{R}^p$ the linear model's parameters

### Exercise 3

(a) State the loss function for ordinary least squares (OLS)

(b) Formulate OLS as a minimization problem

(c) Compute the solution $\hat{\beta}^{OLS}$

(d) What happens if $X^T X$ is singular (not invertible)?

### Exercise 4 - Bias and Variance of OLS

slide 16

### Exercise 5 - Optimality of OLS

slide 17

## Ridge Regression

### Exercise 6 - Bias and Variance

slide 33

### Exercise 7 - Performance

slide 34

## Ridge Logistic Regression

### Exercise 8

slide 46