CS 189 Fall 2017 Introduction to Machine Learning DIS1

1 Unitary invariance

Prove that the regular Euclidean norm (also called the 2-norm) is unitary invariant; in other words, the 2-norm of a vector is the same, regardless of how you apply a rigid transformation to the vector (i.e., rotate or reflect). Note that rigid transformation of a vector $\vec{v} \in \mathbb{R}^d$ means multiplying by an orthogonal $U \in \mathbb{R}^{d \times d}$.

2 Eigenvalues

(a) Let A be an invertible matrix. Show that if $\vec{v}$ is an eigenvector of A with eigenvalue λ, then it is also an eigenvector of A^{-1} with eigenvalue λ^{-1}.

(b) A square and symmetric matrix A is said to be positive semidefinite (PSD) ($A \succeq 0$) if $\forall \vec{v} \neq 0, \vec{v}^T A \vec{v} \geq 0$. Show that A is PSD if and only if all of its eigenvalues are nonnegative.

Hint: Use the eigendecomposition of the matrix A.

3 Least Squares (using vector calculus)

1. In ordinary least-squares linear regression, there is typically no $\vec{x}$ such that $A\vec{x} = \vec{y}$ (these are typically overdetermined systems — too many equations given the number of unknowns). Hence, we need to find an approximate solution to this problem. The residual vector will be $\vec{r} = A\vec{x} - \vec{y}$ and we want to make it as small as possible. The most common case is to measure the residual error with the standard Euclidean 2-norm. So the problem becomes:

$$\min_{\vec{x}} \|A\vec{x} - \vec{y}\|_2^2$$

Where $A \in \mathbb{R}^{m \times n}, \vec{x} \in \mathbb{R}^n, \vec{y} \in \mathbb{R}^m$. Derive using vector calculus an expression for an optimal estimate for $\vec{x}$ for this problem assuming A is full rank.

2. What should we do if A is not full rank?