

# Statistical Methods of Machine Learning

## Assignment 1

Martin Grunbaum  
Martin Jorgensen  
Thomas Barnholdt

February 6, 2014

### I.1.1.1

Given

$$a = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Then  $a^T b = 9$

### I.1.1.2

The  $l_2$ -norm or *Euclidean norm*  $\|a\| = \sqrt{1^2 + 2^2 + 2^2} = 3$

### I.1.1.3

The outer product

$$ab^T = \begin{bmatrix} 3 & 2 & 1 \\ 6 & 4 & 2 \\ 6 & 4 & 2 \end{bmatrix}$$

### I.1.1.4

The inverse matrix of  $M$  is

$$M^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$$

### I.1.1.5

The matrix-vector product  $Ma = \begin{pmatrix} 1 \\ 8 \\ 4 \end{pmatrix}$

### I.1.1.6

$$A = ab^T = \begin{bmatrix} 3 & 6 & 6 \\ 2 & 4 & 4 \\ 1 & 2 & 2 \end{bmatrix}$$

### I.1.1.7

The rank of  $A = 1$ , because the rows are linearly dependent. We can verify this by observing that the third row can produce the first and second rows with a multiple, e.g. the first row (3 6 6) is the same as the third row (1 2 2) x 3.

### I.1.1.8

As  $A$  is not full rank, it is not invertible.

### I.1.2.1

The derivative of  $f(w) = (wx + b)^2$  with respect to  $w$  is

$$(wx + b)^2 = w^2x^2 + 2wxb + b^2 = 2x^2w + 2xb = 2x(xw + b)$$

### I.1.2.2

In general

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2}$$

Therefore, differentiating for  $w$  we get:

$$\begin{aligned}f(x) &= 1 \\f'(x) &= 0 \\g(x) &= (wx + b)^2 \\g'(x) &= 2x(wx + b) \\ \left(\frac{f}{g}\right)'(w) &= \frac{0 \cdot (wx + b)^2 - 1 \cdot 2x(wx + b)}{((wx + b)^2)^2} \\ &= \frac{-1 \cdot 2x(wx + b)}{(wx + b)^4} \\ &= \frac{-2x}{(wx + b)^3}\end{aligned}$$

### I.1.2.3

In general

$$(f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

Therefore, differentiating for  $x$  we get:

$$\begin{aligned}f(x) &= x \\g(x) &= e^x \\(f \cdot g)'(x) &= 1e^x + xe^x\end{aligned}$$