Machine Learning Exercise Sheet 1

Math Refresher

Anastasia Stamatouli 03710902

Martin Meinel : 03710370

20/10/2019

• Linear Algebra

Problem ① :

. $A \in \mathbb{R}^{m \times n}$

$B \in \mathbb{R}^{m}$

$C \in \mathbb{R}^{n \times p}$

$D \in \mathbb{R}^{q}$

$E \in \mathbb{R}^{n \times m}$

$F \in \mathbb{R}$

Problem ② :

$$f(x) = x^T M^T x$$

Problem ③ :

a) • In order a solution to exist, b should belong to the column space of A.

- In order a solution to be unique the number of columns n should be equal to rank (full rank)

- In order a solution to be unique the determinant $\det(A) \neq 0$.

b) We can calculate the determinant by using the eigenvalues. The product of the eigenvalues. is the same as the determinant A.

In our case: $\det(A) = -5 * 0 * 1 * 1 * 3 = 0$, so there is no unique solution.

Problem 4 :

$B = A^{-1}$

$Ax = \lambda x \Leftrightarrow A^{-1}Ax = \lambda A^{-1}x \Leftrightarrow Ix = \lambda A^{-1} \cdot x$

$\lambda \neq 0$ since A invertible

$\Longleftrightarrow \quad A^{-1}x = \frac{1}{\lambda}x \Leftrightarrow Bx = \frac{1}{\lambda}x$

So if the eigenvalue of $A^{-1} = \frac{1}{\lambda}$ then the eigenvalue of $A = \lambda$

Problem 5 :

Let's say λ eigenvalue of M. Then there exist eigenvector $v \in V$ such that

$$Ax = \lambda x$$

$$0 \leq x^T A x = \lambda x^T x$$

Since $x^T x$ is positive for all x $(\|x\|_2^2)$, it implies that λ is non-negative

Problem 6 :

Since $B = A^T A$ is positive semi-definite we can write it as follows:

$$x^T A^T A x = (Ax)^T (Ax) = \|Ax\|_2^2 \geq 0$$

Therefore, the matrix B is positive semi-definite for any choice of A.

• Calculus

Problem (7) :

a) i) unique solution $\rightarrow$ $a > 0$

ii) infinitely many solution $\rightarrow$ $a, b = 0$

iii) no solution , $a \leq 0$, $b \neq 0$

b) $f'(x) = 0 \Leftrightarrow ax + b = 0 \Leftrightarrow x = \frac{-b}{a}$ (there is a critical point at $x = \frac{-b}{a}$)

$f''(x) = (ax+b)'' = a$

if $a > 0$ then $f\left(-\frac{b}{a}\right)$ minimum

Problem 8:

a) $\nabla g(x) = \left(\frac{1}{2}x^T A x + b^T x + c \right)' = Ax + b$

$\nabla^2 g(x) = A$

In order to have a unique solution, A has to be positive definite matrix

b) if the matrix has negative eigenvalue then we lose the PSD. which means that there might not be unique solution. That's why we need the matrix to be PSD.

c) $\nabla g(x) = 0 \Leftrightarrow Ax + b = 0 \Leftrightarrow Ax = -b \Leftrightarrow$
$A^{-1}Ax = -A^{-1}b \Leftrightarrow \boxed{x = -A^{-1}b}$

and $Hf(x^*)$ has to be positive definite.

$Hf(x^*) = A$ which is PD.

Scanned with CamScanner

• Probability Theory

Problem 10:

$$p(A|B,C) = \frac{p(A,B,C)}{p(B|C)\cdot p(C)} = \frac{p(c|a,b)\,p(a|b)\,p(b)}{p(b|c)\cdot p(c)}$$

$$= \frac{p(c|a,b)\cdot p(a)\cdot p(b)}{p(b|c)\cdot p(c)} = \frac{p(a,b|c)\cdot \cancel{p(c)}\cdot \cancel{p(a)\,p(b)}}{p(b|c)\,\cancel{p(c)}\,p(\cancel{a,b})}$$

$$= \frac{p(a,b|c)}{p(b|c)} = \frac{p(a,b,c)}{p(c)\,p(b|c)} = \frac{p(b|a,c)\cdot p(a,c)}{p(c)\cdot p(b|c)}$$

$$= \frac{p(b|a,c)\cdot p(a,c)}{p(b|c)}$$

We cannot say that $p(b|a,c) = p(b|c)$. They are equal only if they are independent. But we don't know if they are. So this statement does not hold.

Problem 11 :

1) $$p(a) = \int\int p(a,b,c)\, db\, dc$$

2) $$p(c|a,b) = \frac{p(c,a,b)}{\int p(a,b,c)\, dc} = \frac{p(a,b,c)}{\int p(a,b,c)\, dc}$$

3) $$p(b|c) = \frac{p(b|c)}{p(c)} = \frac{\int p(b,c,a)\, da}{p(c)}$$

$$= \frac{\int p(b,c,a)\, da}{\int\int p(c,b,a)\, db\, da}$$

Problem 12 :

ND : " you don't have the disease"

PD : " you have the disease"

PT : " positive test"

NT : " negative test"

$P(ND) = 1 - 0.001 = 0.999$

$$P(PT) = P(PT, PD) + P(PT, ND) =$$

$$P(PT|PD) \cdot P(PD) + P(PT|ND) \cdot P(ND) =$$

$$0.00095 + 0.05 * 0.999 = 0.0509$$

$$P(PD|PT) = \frac{P(PT|PD) \cdot P(PD)}{P(PT)} = \frac{0.95 * 0.001}{0.0509}$$

$$= 0.01866405$$

Problem 13 :

$$E(ax + bx^2 + c) = E(ax) + E(bx^2) + E(c)$$

$$= aE(x) + bE(x^2) + E(c) \overset{(1)}{=}$$

$$a \cdot \mu + b\sigma^2 + b\mu^2 + c$$

$$Var(x) = \sigma^2 \Leftrightarrow E[x^2] - [E[x]]^2 = \sigma^2 \Leftrightarrow$$

$$E[x^2] = \sigma^2 + \mu^2 \quad (1)$$

Problem 14 :

1) $E(g(x)) = E(Ax) = AE(x) = A\mu$

2) $E[g(x)\, g(x)^T] = E(Ax(Ax)^T) =$
$A\, E(xx^T)\, A^T$

3) $E(g(x)^T g(x)) = E((Ax)^T Ax) = E(\|Ax\|_2^2)$

- Let's set $Y = Ax$, then $Y \sim N(\mu, A\Sigma A^T)$
- $Var(y) = A\Sigma A^T \Leftrightarrow E(Y_i^2) - (E[y_i])^2 = A\Sigma A^T$
$\Leftrightarrow E(y_i^2) = A\Sigma A^T + \mu^2 \quad (2)$

$E(\|Ax\|_2^2) = E(Y^T Y) = \sum_{i=1}^{n} E Y_i^2 = \sum_{i=1}^{n} (A\Sigma A^T + \mu^2)$

$= \sum_{i=1}^{n} A\Sigma A^T + \sum_{i=1}^{n} \mu^2 = n(A\Sigma A^T + \mu^2)$

4) X random vector with covariance matrix $\Sigma \in \mathbb{R}^{n\times n}$

$Cov[Ax] = E[(Ax - E[Ax])(Ax - E[Ax])^T] =$
$= E[(Ax - AE[x])(Ax - AE[x]^T)] =$
$= E[A(x - E[x])(x - E[x]^T A^T] =$
$= A E[(x - E[x])(x - E[x]^T] A^T =$
$A\Sigma A^T$

Scanned with CamScanner