COT 5615 Math for Intelligent Systems, Fall 2021

Midterm 2

1. (20 pts)

   (a) It can be expressed as n equations by squaring both sides and expand the norm-squared terms as

$$\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{y}_i^T\boldsymbol{x} + \boldsymbol{y}_i^T\boldsymbol{y}_i = \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{y}_{n+1}^T\boldsymbol{x} + \boldsymbol{y}_{n+1}^T\boldsymbol{y}_{n+1}, \quad i = 1, 2, \ldots, n.$$

   Then we have

$$\begin{bmatrix} (\boldsymbol{y}_1 - \boldsymbol{y}_{n+1})^T \\ (\boldsymbol{y}_2 - \boldsymbol{y}_{n+1})^T \\ \vdots \\ (\boldsymbol{y}_n - \boldsymbol{y}_{n+1})^T \end{bmatrix}_{n\times n} \boldsymbol{x} = \frac{1}{2}\begin{bmatrix} \|\boldsymbol{y}_1\|^2 - \|\boldsymbol{y}_{n+1}\|^2 \\ \|\boldsymbol{y}_2\|^2 - \|\boldsymbol{y}_{n+1}\|^2 \\ \vdots \\ \|\boldsymbol{y}_n\|^2 - \|\boldsymbol{y}_{n+1}\|^2 \end{bmatrix}_n .$$

   Therefore, $\boldsymbol{A} = \begin{bmatrix} (\boldsymbol{y}_1 - \boldsymbol{y}_{n+1})^T \\ (\boldsymbol{y}_2 - \boldsymbol{y}_{n+1})^T \\ \vdots \\ (\boldsymbol{y}_n - \boldsymbol{y}_{n+1})^T \end{bmatrix}_{n\times n}$, $\boldsymbol{b} = \frac{1}{2}\begin{bmatrix} \|\boldsymbol{y}_1\|^2 - \|\boldsymbol{y}_{n+1}\|^2 \\ \|\boldsymbol{y}_2\|^2 - \|\boldsymbol{y}_{n+1}\|^2 \\ \vdots \\ \|\boldsymbol{y}_n\|^2 - \|\boldsymbol{y}_{n+1}\|^2 \end{bmatrix}_n$.

   (b) Define $\boldsymbol{Y} = [\ \boldsymbol{y}_1 \ \cdots \ \boldsymbol{y}_n\ ]$, then $\boldsymbol{Ax} = \boldsymbol{b}$ can be expressed as

$$\boldsymbol{Y}^\top\boldsymbol{x} - \boldsymbol{1}\boldsymbol{y}_{n+1}^\top\boldsymbol{x} = \boldsymbol{b}.$$

   If we introduce a new variable $z = -\boldsymbol{y}_{n+1}^\top\boldsymbol{x}$, then $\boldsymbol{Ax} = \boldsymbol{b}$ can be expressed as

$$\boldsymbol{Y}^\top\boldsymbol{x} + z\boldsymbol{1} = \boldsymbol{b}.$$

   Together with $z = -\boldsymbol{y}_{n+1}^\top\boldsymbol{x}$, they form a set of $n+1$ equations with $n+1$ unknowns:

$$\begin{bmatrix} \boldsymbol{Y}^\top & \boldsymbol{1} \\ \boldsymbol{y}_{n+1}^\top & 1 \end{bmatrix}\begin{bmatrix} \boldsymbol{x} \\ z \end{bmatrix} = \begin{bmatrix} \boldsymbol{b} \\ 0 \end{bmatrix}.$$

   The solution to $\boldsymbol{Ax} = \boldsymbol{b}$ is unique if and only if the solution to the above equation is unique, in which case the $(n+1)\times(n+1)$ matrix

$$\begin{bmatrix} \boldsymbol{y}_1 & \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_{n+1} \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

   is invertible.

   Sometimes it makes sense to increase the problem dimension. For example, if $\boldsymbol{Y}$ is sparse, the resulting $\boldsymbol{A}$ matrix may not be sparse. By solving the system with $n+1$ variables, the sparsity pattern is kept, and the complexity could be a lot lower by using an appropriate algorithm.

2. (24 pts)

(a) $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \\ 1 & -1 & 1 \\ -1 & -4 & 0 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ -2 \end{bmatrix}$.

(b) $\boldsymbol{A} = \begin{bmatrix} 0 & -6 \\ -4 & 3 \\ 1 & 8 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}$.

(c) $\boldsymbol{A} = \begin{bmatrix} -6\sqrt{2} & 0 \\ -4\sqrt{3} & 3\sqrt{3} \\ 2 & 16 \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} -4\sqrt{2} \\ \sqrt{3} \\ 6 \end{bmatrix}$.

(d) $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{I}_{n\times n} \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{0}_n \end{bmatrix}$.

(e) $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} \\ \sqrt{2}\boldsymbol{F} \end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix} \boldsymbol{d} \\ \sqrt{2}\boldsymbol{g} \end{bmatrix}$.

(f) $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{D}^{\frac{1}{2}} \end{bmatrix}$, where $\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{D}^{\frac{1}{2}} = \boldsymbol{D}$.

$\boldsymbol{b} = \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{0}_n \end{bmatrix}$.

3. (16 pts)

(a) False. In this case $\alpha = v < 0$.

(b) True. We will predict everything negative, so the false positive reate becomes zero.

(c) False. In this case $\alpha = v/2 < 0$.

(d) True. This is equivalent to $\hat{y} = \text{sign}(\boldsymbol{x}^T\boldsymbol{w} + 2v)$, i.e., $\alpha = -v > 0$.

4. (20 pts)

$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

$$\boldsymbol{A}_{7\times 4} = \begin{bmatrix} 1/2 & t_1/2 & t_1^2/2 & t_1^3/2 \\ 1/2 & t_2/2 & t_2^2/2 & t_2^3/2 \\ 1/2 & t_3/2 & t_3^2/2 & t_3^3/2 \\ 1/2 & t_4/2 & t_4^2/2 & t_4^3/2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1/12 & 1/8 \end{bmatrix}, \boldsymbol{b}_7 = \begin{bmatrix} y_1/2 \\ y_2/2 \\ y_3/2 \\ y_4/2 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

5. (20 pts)

(a) $\boldsymbol{A}$ is invertible if and only if $\boldsymbol{a} \neq 0$.

We use the property that a square matrix is invertible if and only if its columns are linearly independent, i.e., $\boldsymbol{Ax} = 0$ implies $\boldsymbol{x} = 0$. We partition $\boldsymbol{x} = (\boldsymbol{y}, z)$ where $\boldsymbol{y}$ is a n-vector and z is a scalar. Then

$$\boldsymbol{Ax} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{a} \\ \boldsymbol{a}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ z \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} + \boldsymbol{a}z \\ \boldsymbol{a}^T\boldsymbol{y} \end{bmatrix}.$$

Therefore $\boldsymbol{Ax} = 0$ is equivalent to

$$\boldsymbol{y} + \boldsymbol{a}z = 0, \quad \boldsymbol{a}^T\boldsymbol{y} = 0.$$

If $\boldsymbol{a} = 0$, a nonzero $\boldsymbol{x}$ that satisfied this equation is with $\boldsymbol{y} = 0$ and $z = 1$. Therefore $\boldsymbol{A}$ is not invertible if $\boldsymbol{a} = 0$. Suppose $\boldsymbol{a} \neq 0$, then substituting $\boldsymbol{y} = -\boldsymbol{a}z$ in the second equation gives $-z\|\boldsymbol{a}\|^2 = 0$. Since $\boldsymbol{a} \neq 0$, this implies $z = 0$. Then $\boldsymbol{y} = -\boldsymbol{a}z = 0$. This shows that $\boldsymbol{A}$ is invertible.

(b) The inverse is given by

$$\boldsymbol{A}^{-1} = \frac{1}{\|\boldsymbol{a}\|^2} \begin{bmatrix} \|\boldsymbol{a}\|^2\boldsymbol{I} - \boldsymbol{a}\boldsymbol{a}^T & \boldsymbol{a} \\ \boldsymbol{a}^T & -1 \end{bmatrix}.$$

You can directly verify that their product is the $(n+1) \times (n+1)$ identity matrix.

Let's try to derive this result. The ith column of $\boldsymbol{A}^{-1}$ is the solution to the equation $\boldsymbol{Ax} = \boldsymbol{e}_i$.

For the first n column, it can be partitioned as

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{a} \\ \boldsymbol{a}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ z \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} + \boldsymbol{a}z \\ \boldsymbol{a}^T\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_i \\ 0 \end{bmatrix}.$$

Multiply the first block with $\boldsymbol{a}^T$ we get

$$\boldsymbol{a}^T(\boldsymbol{y} + \boldsymbol{a}z) = \boldsymbol{a}^T\boldsymbol{e}_i \Rightarrow z\|\boldsymbol{a}\|^2 = a_i.$$

So $z = a_i/\|\boldsymbol{a}\|^2$ and $\boldsymbol{y} = \boldsymbol{e}_i - (a_i/\|\boldsymbol{a}\|^2)\boldsymbol{a}$. The last column gives us

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{a} \\ \boldsymbol{a}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ z \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} + \boldsymbol{a}z \\ \boldsymbol{a}^T\boldsymbol{y} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Substituting $\boldsymbol{y} = -z\boldsymbol{a}$ into $\boldsymbol{a}^T\boldsymbol{y} = 1$ gives us $z = -1/\|\boldsymbol{a}\|^2$, and so $\boldsymbol{y} = \boldsymbol{a}/\|\boldsymbol{a}\|^2$. Stacking them together

$$\begin{aligned}
\boldsymbol{A}^{-1} &= \begin{bmatrix} \boldsymbol{e}_1 - (a_1/\|\boldsymbol{a}\|^2)\boldsymbol{a} & \cdots & \boldsymbol{e}_n - (a_n/\|\boldsymbol{a}\|^2)\boldsymbol{a} & \boldsymbol{a}/\|\boldsymbol{a}\|^2 \\ a_1/\|\boldsymbol{a}\|^2 & \cdots & a_n/\|\boldsymbol{a}\|^2 & -1/\|\boldsymbol{a}\|^2 \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{I} & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{\|\boldsymbol{a}\|^2} \begin{bmatrix} a_1\boldsymbol{a} & \cdots & a_n\boldsymbol{a} & -\boldsymbol{a} \\ -a_1 & \cdots & -a_n & 1 \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{I} & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{\|\boldsymbol{a}\|^2} \begin{bmatrix} \boldsymbol{a}\boldsymbol{a}^T & -\boldsymbol{a} \\ -\boldsymbol{a}^T & 1 \end{bmatrix} \\
&= \frac{1}{\|\boldsymbol{a}\|^2} \begin{bmatrix} \|\boldsymbol{a}\|^2\boldsymbol{I} - \boldsymbol{a}\boldsymbol{a}^T & \boldsymbol{a} \\ \boldsymbol{a}^T & -1 \end{bmatrix}.
\end{aligned}$$