Final

E4.54
CS4.1

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2010

MSc and EEE/ISE PART IV: MEng and ACGI

PREDICTIVE CONTROL

Monday, 17 May 10:00 am

Time allowed: 3:00 hours

There are SIX questions on this paper.

Answer FOUR questions.

All questions carry equal marks.

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker(s) : E.C. Kerrigan

Second Marker(s) : S. Evangelou

PREDICTIVE CONTROL

1. Consider the following finite-horizon discrete-time optimal control problem:

$$\min_{u_0,u_1,\ldots,u_{N-1}} \sum_{k=0}^{N-1} (\|Qx_k\|_1 + \|Ru_k\|_1)$$

where the system dynamics are given by

$$x_{k+1} = Ax_k + Bu_k, \quad k = 0,1,\ldots,N-1,$$

where the states $x_k \in \mathbb{R}^n$, inputs $u_k \in \mathbb{R}^m$ and weighting matrices $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$.

a) Show that the above problem can be solved by setting up and solving a linear program (LP) of the form

$$\min_{\theta} h^T\theta$$

subject to the constraints

$$L\theta \leq s.$$

In other words, derive expressions for θ, h, s and L such that the solution of the optimal control problem is easily found from the solution of the LP. [16]

b) What are the sizes of the vectors θ, h, s and the matrix L in terms of N, m, n and p? [4]

2. Consider the following discrete-time system:

$$x_{k+1} = Ax_k + B(u_k + d_k),$$
$$y_k = Cx_k,$$

where the states $x_k \in \mathbb{R}^n$, inputs $u_k \in \mathbb{R}^m$, measured outputs $y_k \in \mathbb{R}^p$ and an unknown, constant input disturbance $d_k \in \mathbb{R}^m$, i.e.

$$d_{k+1} = d_k \text{ for } k = 0, 1, \ldots$$

a) Show that one can construct a stable observer to estimate the unknown disturbance if and only if (C, A) is detectable and

$$\begin{pmatrix} A - I & B \\ C & 0 \end{pmatrix}$$

is full column rank. [8]

b) Explain why, for the system above, it is possible to construct a stable observer to estimate the unknown disturbance only if the number of measured outputs p is more than or equal to the number of control inputs m. [2]

c) Suppose now that $m = p$ and that you have constructed a stable observer to provide an estimate of the state $\hat{x}_k$ and disturbance $\hat{d}_k$ at each time instant. Show that the control law

$$u_k = K(\hat{x}_k - x_\infty) + u_\infty,$$

where the control gain matrix $K \in \mathbb{R}^{m \times n}$ is such that all eigenvalues of $A + BK$ have magnitude less than one and (x_∞, u_∞) satisfies

$$\begin{pmatrix} x_\infty \\ u_\infty \end{pmatrix} = \begin{pmatrix} A - I & B \\ C & 0 \end{pmatrix}^{-1} \begin{pmatrix} -B\hat{d}_k \\ r \end{pmatrix},$$

will ensure that

$$y_\infty := \lim_{k \to \infty} y_k = r.$$

[10]

3. We are interested in solving the following optimal control problem :

$$\min_{u_0,\ldots,u_{N-1}} \sum_{k=0}^{N-1} \|Qx_k + Ru_k\|_2^2$$

subject to the constraints

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \quad k = 0,1,\ldots,N-1 \\ Fx_k + Gu_k &\leq g, \quad k = 0,1,\ldots,N-1 \end{aligned}$$

where the states $x_k \in \mathbb{R}^n$, inputs $u_k \in \mathbb{R}^m$, $F \in \mathbb{R}^{p\times n}$, $G \in \mathbb{R}^{p\times m}$, $P \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{q\times n}$, $R \in \mathbb{R}^{q\times m}$ and $g \in \mathbb{R}^p$.

a) Show that the above problem can be solved by setting up and solving a quadratic program (QP) of the form

$$\min_{\theta} \ \theta^T H\theta + h^T\theta$$

subject to the constraints

$$L\theta \leq s.$$

In other words, derive expressions for θ, h, s, H and L such that the solution of the optimal control problem is easily found from the solution of the QP. [15]

b) What are the sizes of θ, h, s, H and L and in terms of N, m, n, p and q? [5]

4. a) List a number of applications in industry where predictive control has been successfully applied. Discuss why predictive control has been so successful in these applications. [8]

b) List a number of applications in industry where predictive control has not yet been applied or has had limited success. Discuss some of the reasons why predictive control has not yet been applied to or has had limited success in these applications. [8]

c) Discuss some of the technological and theoretical problems that need to be solved before predictive control can be applied more widely in the applications listed in part b). [4]

5. a) Consider the optimization problem

$$\varepsilon^* := \min_{(\varepsilon,\theta)} \varepsilon$$

subject to the constraints

$$M\theta \le b + \mathbf{1}_p \varepsilon,$$

where $\mathbf{1}_p$ is a column vector of ones of length p, $M \in \mathbb{R}^{p\times q}$, $\theta \in \mathbb{R}^q$, $b \in \mathbb{R}^p$ and $\varepsilon \in \mathbb{R}$.

Show that the set $\{\theta \mid M\theta \le b\}$ is non-empty if and only if $\varepsilon^* \le 0$. [8]

b) For a given initial state x_0, we are interested in determining whether an input sequence $(u_0, u_1, \ldots, u_{N-1})$ exists that satisfies the following constraints:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \quad k = 0,1,\ldots,N-1, \\ g \le Fx_k &+ Gu_k \le h, \quad k = 0,1,\ldots,N-1, \end{aligned}$$

where the states $x_k \in \mathbb{R}^n$, inputs $u_k \in \mathbb{R}^m$, vectors $g \in \mathbb{R}^r$, $h \in \mathbb{R}^r$ and the matrices are all of compatible size.

Show that we can establish whether a feasible input sequence exists by solving a suitably-defined optimization problem of the form considered in part a). In other words, derive suitable expressions for θ, M and b. [10]

c) What are the sizes of θ and b in terms of N, m, n and r? [2]

6. a) Recall the linear least squares problem:

$$\theta^* := \arg\min_{\theta} \|M\theta - b\|_2^2,$$

where $M \in \mathbb{R}^{p\times q}$, $\theta \in \mathbb{R}^q$ and $b \in \mathbb{R}^p$.

Show that the solution θ^* to the above least squares problem is given by the solution to the set of linear equations

$$M^T M\theta^* = M^T b,$$

which is also known as the normal equations. [6]

b) Give a sufficient condition on M that would guarantee the solution to the above least squares problem is unique and given by

$$\theta^* = (M^T M)^{-1} M^T b.$$

Justify your answer. [2]

c) We are interested in solving the following optimal control problem :

$$\left(u_0^*(x_0), \ldots, u_{N-1}^*(x_0)\right) := \arg\min_{(u_0,\ldots,u_{N-1})} \|Px_N\|_2^2 + \sum_{k=0}^{N-1} \left(\|Qx_k\|_2^2 + \|Ru_k\|_2^2\right),$$

where the system dynamics are given by

$$x_{k+1} = Ax_k + Bu_k, \quad k = 0, 1, \ldots, N-1,$$

the states $x_k \in \mathbb{R}^n$, inputs $u_k \in \mathbb{R}^m$ and the weights $P \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$.

Show that the above optimal control problem can be converted into an equivalent least squares problem. In other words, derive expressions for θ, M and b such that the solution of the optimal control problem is easily found from the solution of the least squares problem. [10]

d) Give a sufficient condition on R that would guarantee the solution to the optimal control problem is unique. Justify your answer. [2]

Question 1 PREDICTIVE CONTROL SOLUTIONS 2010 E4.54 CS4.1

(a) Application of theory to a new problem. 1/6

$$\sum_{k=0}^{N-1} \left(\|Qx_k\|_1 + \|Ru_k\|_1 \right) = \left\| \begin{pmatrix} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \end{pmatrix} \right\|_1 \text{, where } \bar{x} := \begin{pmatrix} x_0 \\ \vdots \\ x_{N-1} \end{pmatrix}$$

$$\bar{u} := \begin{pmatrix} u_0 \\ \vdots \\ u_{N-1} \end{pmatrix}, \quad \bar{Q} := I_N \otimes Q, \quad \bar{R} := I_N \otimes R \qquad \bar{u} \in \mathbb{R}^{Nm}, \ \bar{x} \in \mathbb{R}^{Nn}$$

$$\bar{x} = \underbrace{\begin{pmatrix} I \\ A \\ \vdots \\ A^{N-1} \end{pmatrix}}_{\Phi} x_0 + \underbrace{\begin{pmatrix} 0 & 0 & \cdots & & 0 \\ B & 0 & & & \vdots \\ AB & B & \ddots & & 0 \\ \vdots & & & & \\ A^{N-2}B & A^{N-3}B & \cdots & & B \end{pmatrix}}_{\Gamma} \bar{u}$$

$$\Rightarrow \left\| \begin{pmatrix} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \end{pmatrix} \right\|_1 = \left\| \begin{pmatrix} \bar{Q}\Phi x_0 + \bar{Q}\Gamma\bar{u} \\ \bar{R}\bar{u} \end{pmatrix} \right\|_1$$

$$\min_{\bar{u}} \left\| \begin{pmatrix} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \end{pmatrix} \right\|_1 = \min_{\varepsilon, \bar{u}} \mathbb{1}^T \varepsilon \text{ subject to } -\varepsilon \le \begin{pmatrix} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \end{pmatrix} \le \varepsilon, \ \varepsilon \ge 0$$

where $\varepsilon \in \mathbb{R}^{N(n+m)}$ and $\mathbb{1} := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{N(n+m)}$

$$\begin{matrix} -\varepsilon \le \begin{pmatrix} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \end{pmatrix} \le \varepsilon \\ \varepsilon \ge 0 \end{matrix} \iff \left(\begin{array}{c} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \\ \hline -\bar{Q}\bar{x} \\ -\bar{R}\bar{u} \end{array} \right) \le \begin{pmatrix} \varepsilon \\ \varepsilon \end{pmatrix}, \quad -\varepsilon \le 0$$

$$\iff \left(\begin{array}{c} \bar{Q}\Phi x_0 + \bar{Q}\Gamma\bar{u} \\ \bar{R}\bar{u} \\ \hline -\bar{Q}\Phi x_0 - \bar{Q}\Gamma\bar{u} \\ -\bar{R}\bar{u} \end{array} \right) - \left(\begin{array}{c} I \\ \hline I \end{array} \right) \varepsilon \le 0, \quad -I\varepsilon \le 0 \qquad h := \begin{pmatrix} 0 \\ \mathbb{1} \end{pmatrix}$$

$$\iff \underbrace{\left(\begin{array}{c|c} \bar{Q}\Gamma & -I \\ \bar{R} & \\ \hline -\bar{Q}\Gamma & -I \\ -\bar{R} & \\ \hline 0 & -I \end{array} \right)}_{L \in \mathbb{R}^{3N(m+n) \times N(2m+n)}} \underbrace{\begin{pmatrix} \bar{u} \\ \varepsilon \end{pmatrix}}_{\Theta} \le \underbrace{\left(\begin{array}{c} -\bar{Q}\Phi x_0 \\ 0 \\ \hline \bar{Q}\Phi x_0 \\ 0 \\ \hline 0 \end{array} \right)}_{s}$$

(b) Θ is a vector of length $Nm + N(n+m)$

h is a vector of length $N(2m+n)$

L is a matrix of size $3N(n+m) \times N(2m+n)$

s is a vector of length $3N(n+m)$

Question 2

(a) Application of theory to a specific problem.

Construct the augmented system:

$$\begin{pmatrix} x_{k+1} \\ d_{k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} A & B \\ 0 & I \end{pmatrix}}_{\tilde{A}} \begin{pmatrix} x_k \\ d_k \end{pmatrix} + \begin{pmatrix} B \\ 0 \end{pmatrix} u_k.$$

$$y_k = \underbrace{\begin{pmatrix} C & 0 \end{pmatrix}}_{\tilde{C}} \begin{pmatrix} x_k \\ d_k \end{pmatrix}$$

Augmented system detectable $\Leftrightarrow$ $(\tilde{C}, \tilde{A})$ detectable.

PBH test: $\Leftrightarrow \begin{pmatrix} A-\lambda I & B \\ 0 & I-\lambda I \\ C & 0 \end{pmatrix}$ full column rank $\forall |\lambda| \geq 1$

e/values of $\tilde{A}$ = e/values of A and $\{1, 1, \ldots 1\}$

From ~~the~~ PBH test first set of columns linearly independent $\Leftrightarrow$ (C, A) detectable.

The second set of columns linearly independent of the first set of columns except possibly for $\lambda = 1$

$\Leftrightarrow \begin{pmatrix} A-I & B \\ 0 & 0 \\ C & 0 \end{pmatrix}$ full colum rank $\Leftrightarrow \begin{pmatrix} A-I & B \\ C & 0 \end{pmatrix}$ full column rank.

(b) $A-I$ is square so $\begin{pmatrix} A-I & B \\ C & 0 \end{pmatrix}$ ~~has to have~~ is full rank only if number of columns of B $\leq$ ~~as~~ number of rows of C

$\Rightarrow m \leq p$

(c) At equilibrium $x_\infty = A x_\infty + B u_\infty + B d_\infty$

where we want $C x_\infty = r$

$$\Leftrightarrow \begin{pmatrix} A-I & +B \\ C & 0 \end{pmatrix} \begin{pmatrix} x_\infty \\ u_\infty \end{pmatrix} = \begin{pmatrix} -B d_\infty \\ r \end{pmatrix} \quad (*)$$

Closed-loop is: $x_{k+1} = A x_k + BK x_k - BK x_\infty + B u_\infty + B d_\infty$

$= (A+BK) x_k - BK x_\infty + B u_\infty + B d_\infty$

Since $(A+BK)$ stable, at equilibrium, $x_\infty = (A+BK) x_\infty - BK x_\infty + B u_\infty + B d_\infty$

$\Leftrightarrow (A-I) x_\infty + B u_\infty = -B d_\infty$

If $m = p$ & $\begin{pmatrix} A-I & B \\ C & 0 \end{pmatrix}$ full rank $\Rightarrow \begin{pmatrix} A-I & B \\ C & 0 \end{pmatrix}^{-1}$ exists.

and ~~d~~ since (x_∞, u_∞) satisfies $(*)$, we have $C x_\infty = r$.

Question 3

(a) Application of theory to a new problem.

Constraints: $Fx_k + Gu_k \leq g$, $k = 0, 1, \ldots, N-1$

$$\Leftrightarrow \bar{F}\bar{x} + \bar{G}\bar{u} \leq \bar{g}, \text{ where } \bar{x} := \begin{pmatrix} x_0 \\ \vdots \\ x_{N-1} \end{pmatrix}, \bar{u} := \begin{pmatrix} u_0 \\ \vdots \\ u_{N-1} \end{pmatrix}$$

$$\bar{F} := I_N \otimes F, \quad \bar{G} := I_N \otimes G, \quad \bar{g} := 1_N \otimes g$$

where $1_N := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^N$.

$$\Rightarrow \bar{g} \in \mathbb{R}^{Np}$$

$$\bar{x} = \underbrace{\begin{pmatrix} I \\ A \\ A^2 \\ \vdots \\ A^{N-1} \end{pmatrix}}_{\Phi} x_0 + \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 \\ B & 0 & & \vdots \\ AB & B & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A^{N-2}B & \cdots & & AB\,B \end{pmatrix}}_{\Gamma} \bar{u} = \Phi x_0 + \Gamma \bar{u}$$

$$\Theta := \bar{u}$$

$$\Rightarrow \bar{F}\bar{x} + \bar{G}\bar{u} \leq \bar{g} \Leftrightarrow \underbrace{(\bar{F}\Gamma + \bar{G})}_{L}\underbrace{\bar{u}}_{\Theta} \leq \underbrace{\bar{g} - \Phi x_0}_{S}$$

Cost:

$$\sum_{k=0}^{N-1} \|Qx_k + Ru_k\|_2^2 = \|\bar{Q}\bar{x} + \bar{R}\bar{u}\|_2^2, \text{ where } \begin{array}{l} \bar{Q} := I_N \otimes Q \\ \bar{R} := I_N \otimes R \end{array}.$$

$$= \|\bar{Q}\Phi x_0 + (\bar{Q}\Gamma + \bar{R})\bar{u}\|_2^2$$

$$= x_0^T \Phi^T \bar{Q}^T \bar{Q} \Phi x_0 + \bar{u}^T \underbrace{(\bar{Q}\Gamma + \bar{R})^T(\bar{Q}\Gamma + \bar{R})}_{H} \bar{u} + \underbrace{2x_0^T \Phi^T \bar{Q}^T (\bar{Q}\Gamma + \bar{R})}_{h^T} \bar{u}$$

$$\Rightarrow h := 2(\bar{Q}\Gamma + \bar{R})^T \bar{Q} \Phi x_0$$

(b) Θ is a vector of length Nm

h is a " " "

H " " matrix with Nm rows and Nm columns

L " " " " Np rows and Nm columns

s " " vector " Np rows.

Question 4

Bookwork – Students were given survey & tutorial papers and recommended textbooks to read.

Sample solution: (Other answers clearly possible)

(a) Petrochemical + pulp & paper. Processes relatively slow compared to computation time. Predictive control easy to understand. Optimal operation often close or on constraints & predictive control can handle constraints in a straightforward manner

(b) ~~A~~ Commercial ~~a~~ aeroplanes and automobiles.
Dynamics often fast compared to computation power.
Lack of guarantees of robustness
Industries more conservative.

(c) Faster algorithms for solving the optimization problems ~~easily~~ on-line. Robust formulations. Education of engineers designing controllers. Cost of implementation needs to be reduced.

Question 5

(a) "Bookwork" – done in lectures

I) The set is non-empty $\Leftrightarrow \exists \theta : M\theta \le b \Leftrightarrow \exists \theta : M\theta \le b + 1_p \varepsilon$ with $\varepsilon = 0$

But $\varepsilon^* \le \varepsilon \Rightarrow \varepsilon^* \le 0$

II) $\varepsilon^* \le 0 \Rightarrow \exists\, \varepsilon \le 0$ and $\theta : M\theta \le b + 1_p \varepsilon \le b \Rightarrow$ set non-empty

---

(b) $g \le Fx_k + Gu_k \le h,\ k = 0, 1, \dots, N-1$

$\Leftrightarrow \bar{g} \le \bar{F}\bar{x} + \bar{G}\bar{u} \le \bar{h}$, where $\bar{x} := \begin{pmatrix} x_0 \\ \vdots \\ x_{N-1} \end{pmatrix}$, $\bar{u} := \begin{pmatrix} u_0 \\ \vdots \\ u_{N-1} \end{pmatrix}$

$\bar{F} := I_N \otimes F$, $\bar{G} := I_N \otimes G$, $\bar{g} := 1_N \otimes g$, $\bar{h} := 1_N \otimes h$

$1_N := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^N \Rightarrow \bar{g} \in \mathbb{R}^{Nr}$

$$\bar{x} = \begin{pmatrix} I \\ A \\ A^2 \\ \vdots \\ A^{N-1} \end{pmatrix} x_0 + \begin{pmatrix} 0 & 0 & \cdots & & 0 \\ B & 0 & & & \vdots \\ AB & \ddots & & & \\ \vdots & & & B & 0 \\ A^{N-2}B & \cdots & & AB & B \end{pmatrix} \bar{u} =: \Phi x_0 + \Gamma \underline{u}$$

$$\left.\begin{aligned} \bar{F}\bar{x} + \bar{G}\bar{u} &\le \bar{h} \\ -\bar{F}\bar{x} - \bar{G}\bar{u} &\le -\bar{g} \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} \bar{F}\Phi x_0 + \bar{F}\Gamma\bar{u} + \bar{G}\bar{u} &\le \bar{h} \\ -\bar{F}\Phi x_0 - \bar{F}\Gamma\bar{u} - \bar{G}\bar{u} &\le -\bar{g} \end{aligned}\right.$$

$$\Leftrightarrow \underbrace{\begin{pmatrix} \bar{F}\Gamma + \bar{G} \\ -\bar{F}\Gamma - \bar{G} \end{pmatrix}}_{M} \underbrace{\bar{u}}_{\theta} \le \underbrace{\begin{pmatrix} \bar{h} - \bar{F}\Phi x_0 \\ -\bar{g} + \bar{F}\Phi x_0 \end{pmatrix}}_{b}$$

(c) θ is a vector of length Nm

b " " " " $2Nr$

Question 6

6/6

(a) "Bookwork" - done in lectures.

$$f(\theta) := \|M\theta - b\|_2^2 = b^T b - 2b^T M\theta + \theta^T M^T M \theta$$

$$\nabla_\theta f = 2M^T M\theta - 2M^T b = 0$$

$$\iff M^T M\theta = M^T b$$

$$\nabla_\theta^2 f = 2M^T M \succeq 0 \quad (M^T M = 0 \text{ not possible unless } M = 0)$$

$\Rightarrow$ stationary point is a ~~max~~ minimum.

(b) M full rank $\Rightarrow M^T M > 0 \Rightarrow (M^T M)^{-1}$ exists
$\Rightarrow$ solution unique.

(c)
$$\|Px_N\|_2^2 + \sum_{k=0}^{N-1} \|Qx_k\|_2^2 + \|Ru_k\|_2^2 = \left\| \begin{pmatrix} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \end{pmatrix} \right\|_2^2, \quad \bar{u} := \begin{pmatrix} u_0 \\ \vdots \\ u_{N-1} \end{pmatrix}$$

$$\bar{x} := \begin{pmatrix} x_0 \\ \vdots \\ x_N \end{pmatrix}, \quad \bar{R} := I_N \otimes R, \quad \bar{Q} := \begin{pmatrix} I_N \otimes Q & 0 \\ 0 & P \end{pmatrix} \in \mathbb{R}^{(N+1)n \times (N+1)n}$$

$$\bar{x} = \underbrace{\begin{pmatrix} I \\ A \\ \vdots \\ A^N \end{pmatrix}}_{\Phi} x_0 + \underbrace{\begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ B & 0 & \cdots & \cdots & \vdots \\ \vdots & B & \cdots & \cdots & 0 \\ A^{N-1}B & A^{N-2}B & \cdots & AB & B \end{pmatrix}}_{\Gamma} \bar{u} = \Phi x_0 + \Gamma \bar{u}$$

$$\Rightarrow \begin{pmatrix} \bar{Q}\bar{x} \\ \bar{R}\bar{u} \end{pmatrix} = \begin{pmatrix} \bar{Q}\Phi x_0 + \bar{Q}\Gamma\bar{u} \\ \bar{R}\bar{u} \end{pmatrix} = \underbrace{\begin{pmatrix} \bar{Q}\Gamma \\ \bar{R} \end{pmatrix}}_{M} \underbrace{\bar{u}}_{\theta} + \underbrace{\begin{pmatrix} \bar{Q}\Phi x_0 \\ 0 \end{pmatrix}}_{-b}$$

---

d) R full rank $\Rightarrow \bar{R}$ full rank $\Rightarrow M$ full rank

$\Rightarrow (M^T M) > 0 \Rightarrow (M^T M)^{-1}$ exists $\Rightarrow$ solution unique

(see part b).
