3.1. Given $z \in \mathbb{R}^n$, $g \in \mathbb{R}^n$ find

NumCSE 16
2.M.16
Ex. 3.1
G. Acceputo

$$M^* = \operatorname*{argmin}_{M \in \mathbb{R}^{n \times n},\, Mz = g} \|M\|_F \quad (1).$$

where $\|M\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |m_{ij}|^2}$ is the Frobenius norm of M.

3.1a. Reformulate Eq. (1) to:

$$x^* = \operatorname*{argmin}_{x \in \mathbb{R}^N,\, Cx = d,\, (N \neq n^2)} \|Ax - b\|_2 \quad (2).$$

1.) We now have to first move from the matrix norm $\|\cdot\|_F$ to the vector norm $\|\cdot\|_2$ and define A, x, and b such that

$$\|Ax - b\|_2 = \sqrt{\sum_{i=1} \sum_{j=1} |m_{ij}|^2} \left(= \|M\|_F\right)$$

By defining $m^* = \text{vec}(M^{*T}) \in \mathbb{R}^{n^2}$ (concatenation of the rows of M^*),

we can see that Eq (2) actually becomes

$$m^* = \operatorname*{argmin}_{m \in \mathbb{R}^{n^2},\, Cm = d} \|m\|_2 \,, \quad (3).$$

since $\|m\|_2 = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |m_{ij}|^2}$. Thus, define A, x, and b

s.t.

$$Ax - b = m. \Rightarrow \|Ax - b\|_2 = \|m\|_2.$$

Set $A = I_{n^2}$, $x = m$ and $b = 0$ and we're done.

2.) We now have to find a way to expren $\underbrace{Mz = g}_{\text{linear constraints}}$ in terms of $Cm = d$.

$$Mz = \begin{pmatrix} [\text{---}\; M_{1:} \;\text{---}] \cdot \begin{bmatrix} | \\ z \\ | \end{bmatrix} \\ [\text{---}\; M_{2:} \;\text{---}] \cdot \begin{bmatrix} | \\ z \\ | \end{bmatrix} \\ \vdots \\ [\text{---}\; M_{n:} \;\text{---}] \cdot \begin{bmatrix} | \\ z \\ | \end{bmatrix} \end{pmatrix} = g.$$

$$\Rightarrow \quad m = \operatorname{vec}(M^T) = \begin{bmatrix} \begin{pmatrix} | \\ M_{1:} \\ | \end{pmatrix} \\ \begin{pmatrix} | \\ M_{2:} \\ | \end{pmatrix} \\ \vdots \\ \begin{pmatrix} | \\ M_{n:} \\ | \end{pmatrix} \end{bmatrix} \in \mathbb{R}^{n^2}$$

1st row of M

We now need to define C s.t.

$$Cm = \begin{pmatrix} [\text{---}\; z \;\text{---}] \cdot \begin{bmatrix} | \\ M_{1:} \\ | \end{bmatrix} \\ [\text{---}\; z \;\text{---}] \cdot \begin{bmatrix} | \\ M_{2:} \\ | \end{bmatrix} \\ \vdots \\ [\text{---}\; z \;\text{---}] \cdot \begin{bmatrix} | \\ M_{n:} \\ | \end{bmatrix} \end{pmatrix} = d,$$

Since $d = g$ and $[\text{---}\; M_{i:} \;\text{---}] \cdot \begin{bmatrix} | \\ z \\ | \end{bmatrix} = [\text{---}\; z \;\text{---}] \cdot \begin{bmatrix} | \\ M_{i:} \\ | \end{bmatrix}$

since we are working with real matrices and vectors!

Thus, define

$$C = I_n \otimes z^T = \begin{pmatrix} \overbrace{[-z-]}^{z^T} & 0 & \cdots & 0 \\ 0 & [-z-] & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & & \\ 0 & 0 & & [-z-] \end{pmatrix}$$

resulting in

$$Cm = g \Leftrightarrow \begin{pmatrix} [-z-] & & 0 \\ 0 & \ddots & \\ & & [-z-] \end{pmatrix} \begin{bmatrix} \begin{pmatrix} | \\ M_{1:} \\ | \end{pmatrix} \\ \vdots \\ \begin{pmatrix} | \\ M_{n:} \\ | \end{pmatrix} \end{bmatrix} = \overbrace{\begin{pmatrix} [-z-] \cdot \begin{bmatrix} | \\ M_{1:} \\ | \end{bmatrix} \\ \vdots \\ [-z-] \begin{bmatrix} | \\ M_{n:} \\ | \end{bmatrix} \end{pmatrix}}^{g}$$

3.1.b.

$C \in \mathbb{R}^{n \times n^2} \iff C$ is a non-square matrix

"if and only if"

$\iff C$ has full rank iff

$\mathrm{rank}(C) = \min(m,k) = \min(n, n^2)$.

$\Rightarrow$ If z has at least one non-zero component we get n linearly independent row and column vectors,

i.e. $\mathrm{rank}(C) = n = \min(n, n^2)$

3.1.c.

See Remark 2.3.14 for the Block gaussian elimination!

$$\left[\begin{array}{cc|c} I_{n^2} & C^T & 0 \\ C & 0 & g \end{array}\right]$$

Set $\ell := C I_{n^2}^{-1} = C$, then

$$\left[\begin{array}{cc|c} I_{n^2} & C^T & 0 \\ C - C I_{n^2} & 0 - CC^T & g \end{array}\right] = \left[\begin{array}{cc|c} I_{n^2} & C^T & 0 \\ 0 & -CC^T & g \end{array}\right]$$

$$\Rightarrow \begin{bmatrix} I_{n^2} & C^T \\ 0 & -CC^T \end{bmatrix} \begin{bmatrix} m^* \\ p \end{bmatrix} = \begin{bmatrix} 0 \\ g \end{bmatrix}$$

$\Rightarrow \quad -CC^T \cdot p = g \iff p = -(C^T C)^{-1} g$

$\Rightarrow$ (theorem 3.1.18) $\quad C^T C$ is regular iff C has full rank (see 3.1.b)

3.1.f

Show that: $M = \frac{g \cdot z^T}{\|z\|_2^2}$ is a solution

of $\begin{bmatrix} I_{n \cdot} & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} m^* \\ p \end{bmatrix} = \begin{bmatrix} 0 \\ g \end{bmatrix}$.

proof: From 3.1.c. we know that

$$-CC^T p = g$$

Further, we can see that

$$C \cdot C^T = \begin{pmatrix} z^T & & 0 \\ & z^T & \\ 0 & & \ddots \; z^T \end{pmatrix} \cdot \begin{pmatrix} z & & 0 \\ & z & \\ 0 & & \ddots \; z \end{pmatrix}$$

$$= \begin{pmatrix} z^T \cdot z & & \\ & z^T \cdot z & \\ & & \ddots \; z^T \cdot z \end{pmatrix} = \begin{pmatrix} \|z\|_2^2 & & \\ & \ddots & \\ & & \|z\|_2^2 \end{pmatrix}$$

$$= \|z\|_2^2 \cdot I_n$$

Thus, we ~~have~~ first have

$$-CC^T p = g \iff -\|z\|_2^2 \cdot I_n \cdot p = g$$

$$\iff p = -\|z\|_2^2 \cdot g$$

Next, we have

$$I_{n^2} \cdot m^* + C^T p = 0$$

$$\Leftrightarrow \quad m^* = - C^T p \Leftrightarrow m^* = \|z\|_2^2 \cdot C^T g$$

For $C^T g$ we get

$$C^T g = \underbrace{\begin{pmatrix} \begin{pmatrix} | \\ z \\ | \end{pmatrix} & & & 0 \\ & \begin{pmatrix} | \\ z \\ | \end{pmatrix} & & \\ & & \ddots & \\ 0 & & & \begin{pmatrix} | \\ z \\ | \end{pmatrix} \end{pmatrix}}_{\in \mathbb{R}^{n^2 \times n}} \cdot \underbrace{g}_{\in \mathbb{R}^n} = \begin{pmatrix} \begin{pmatrix} | \\ z \\ | \end{pmatrix} \cdot g_1 \\ \begin{pmatrix} | \\ z \\ | \end{pmatrix} \cdot g_2 \\ \vdots \\ \begin{pmatrix} | \\ z \\ | \end{pmatrix} \cdot g_n \end{pmatrix}$$

From 3.1.a. we know that

$$m^* = vec(M^{*T}),$$

meaning that m^* is the vector we obtain by concatenating the rows of M^*.

By reversing this operation on $m^* = \|z\|_2^2 C^T g$, i.e. we "un-vectorize" m^* we get

$$M^* = \|z\|_2^2 \begin{pmatrix} \overbrace{(- z -)}^{z^T} \cdot g_1 \\ (- z -) \cdot g_2 \\ \vdots \\ (- z -) \cdot g_n \end{pmatrix} = \underbrace{g \otimes z}_{\text{outer product}} \cdot \|z\|_2^2$$

3.1.g. Use <u>Eq. 3.6.4</u> from the Script.

For
$$M^* = \underset{M \in \mathbb{R}^{n \times n}}{\operatorname{argmin}} \left\{ \max_{m \in \mathbb{R}^n} \left[L(M, m) \right] \right\}$$

we get:
$$L(M, m) = \|M\|_F^2 + m^T (Mz - g)$$

3.1.h. $\|X\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} |x_{ij}|^2$

$$\operatorname{grad} \phi(X) = \frac{\partial}{\partial X} \phi(X) =$$
$$= \begin{pmatrix} \frac{\partial \phi}{\partial x_{11}} & \cdots & \frac{\partial \phi}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial \phi}{\partial x_{n1}} & \cdots & \frac{\partial \phi}{\partial x_{nn}} \end{pmatrix} = \begin{pmatrix} \frac{\partial \|X\|_F^2}{\partial x_{11}} & \cdots & \frac{\partial \|X\|_F^2}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial \|X\|_F^2}{\partial x_{n1}} & \cdots & \frac{\partial \|X\|_F^2}{\partial x_{nn}} \end{pmatrix}$$

$$= \begin{pmatrix} 2x_{11} & \cdots & 2x_{1n} \\ \vdots & & \vdots \\ 2x_{n1} & \cdots & 2x_{nn} \end{pmatrix} = 2X$$

3.1.i:

$$\operatorname{grad} L(M,m) = 0 \Leftrightarrow \begin{pmatrix} \frac{\partial}{\partial M} L(M,m) \\ \frac{\partial}{\partial m} L(M,m) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

with. $$\frac{\partial f}{\partial M} = \begin{pmatrix} \frac{\partial f}{\partial m_{11}} & \cdots & \frac{\partial f}{\partial m_{1n}} \\ \vdots & & \\ \frac{\partial f}{\partial m_{n1}} & \cdots & \frac{\partial f}{\partial m_{nn}} \end{pmatrix}$$

$\boxed{m \in \mathbb{R}^n, \; M \in \mathbb{R}^{n \times n}}$

1.) $$\frac{\partial}{\partial M} L(M,m) = \underbrace{\frac{\partial}{\partial M} \|M\|_F^2}_{\substack{= 2M \\ \text{(see 3.1.h)}}} + \frac{\partial}{\partial M} m^T M z - \underbrace{\frac{\partial}{\partial M} m^T g}_{=0}.$$

first row of M

$$\Rightarrow \frac{\partial}{\partial M} m^T M z = \frac{\partial}{\partial M} m^T \cdot \begin{pmatrix} (- M_{1:} -) \begin{pmatrix} | \\ z \\ | \end{pmatrix} \\ \vdots \\ (- M_{n:} -) \begin{pmatrix} | \\ z \\ | \end{pmatrix} \end{pmatrix}$$

$$= \frac{\partial}{\partial M} \Big[\underbrace{m_1 \cdot (- M_{1:} -) \cdot \begin{pmatrix} | \\ z \\ | \end{pmatrix} + m_2 \cdot (- M_{2:} -) \cdot \begin{pmatrix} | \\ z \\ | \end{pmatrix} + \dots + m_n \cdot (- M_{n:} -) \cdot \begin{pmatrix} | \\ z \\ | \end{pmatrix}}_{\boxed{d} :=} \Big]$$

$$\Rightarrow \frac{\partial}{\partial M} \underset{\mathbb{R}}{\underset{\cup}{m^T M z}} = \frac{\partial}{\partial M} \underset{\mathbb{R}}{\underset{\cup}{\boxed{d}}}$$

$$= \begin{pmatrix} \frac{\partial d}{\partial m_{11}} & \cdots & \frac{\partial d}{\partial m_{1n}} \\ \vdots & & \\ \frac{\partial d}{\partial m_{n1}} & \cdots & \frac{\partial d}{\partial m_{nn}} \end{pmatrix} = \begin{pmatrix} m_1 \cdot z_1 & m_1 \cdot z_2 & \cdots & m_1 \cdot z_n \\ m_2 \cdot z_1 & \cdots & & m_2 \cdot z_n \\ \vdots & & & \\ m_n \cdot z_1 & \cdots & & m_n \cdot z_n \end{pmatrix}$$

$$= m \otimes z = \underline{\underline{m z^T}}$$

$$\Rightarrow \frac{\partial L(M,m)}{\partial M} = 2M + mz^T = 0$$

2.) $$\frac{\partial}{\partial m} L(M,m) = \underbrace{\frac{\partial}{\partial m} \|M\|_F^2}_{=0} + \frac{\partial}{\partial m} m^T M z - \frac{\partial}{\partial m} m^T g.$$

$$\Rightarrow \frac{\partial}{\partial m} m^T g = \frac{\partial}{\partial m} [m_1 \cdot g_1 + m_2 g_2 + \dots + m_n \cdot g_n].$$

$$\Rightarrow \frac{\partial}{\partial m} m^T g = \begin{pmatrix} \frac{\partial}{\partial m_1} m^T g \\ \vdots \\ \frac{\partial}{\partial m_n} m^T g \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix} = g$$

$$\Rightarrow \frac{\partial}{\partial m} L(M,m) = Mz - g = 0$$

3.A.j.

1.) $\frac{\partial L}{\partial M} = 2M + mz^T = 0$

$$\Leftrightarrow \quad M = -\frac{1}{2} mz^T.$$

2.) $\frac{\partial L}{\partial m} = Mz - g = 0$

$$\Rightarrow (\text{use } M = -\tfrac{1}{2} mz^T) \quad -\frac{1}{2} m \underbrace{z^T z}_{= \|z\|_2^2} - g = 0$$

$$\Leftrightarrow \quad -\frac{1}{2} m \|z\|_2^2 - g = 0$$

$$\Leftrightarrow \quad m = -2 \cdot \frac{g}{\|z\|_2^2}$$

3.) $M = -\frac{1}{2} mz^T \Leftrightarrow (\text{use } m = \frac{-2g}{\|z\|_2^2}) \quad M^* = \frac{mz^T}{\|z\|_2^2}$