Project in Optimiziation Theory - TMA 4180

April 4, 2014

Abstract

A truss is a mechanical structure constructed of straight elastic members connected at joints. In this project we have used optimization theory to maximize the stiffness of the structure.

Equations for questions 1-3

$$\begin{aligned} \underset{(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}})}{\text{minimize}} \quad & \frac{1}{2}\sum_{j=1}^{m} \frac{\ell_j q_j^2}{E_j A_j} = g(\boldsymbol{q}) \\ \text{subject to} \quad & \boldsymbol{B}\boldsymbol{q} = \boldsymbol{I}_{\text{supp}}\boldsymbol{f}^{\text{supp}} + \boldsymbol{I}_{\text{ext}}\boldsymbol{f}^{\text{ext}} \end{aligned} \tag{1}$$

$$\begin{aligned} \boldsymbol{D}\boldsymbol{q} &= \boldsymbol{B}^T\boldsymbol{u} \\ \boldsymbol{B}\boldsymbol{q} &= \boldsymbol{I}_{\text{supp}}\boldsymbol{f}^{\text{supp}} + \boldsymbol{I}_{\text{ext}}\boldsymbol{f}^{\text{ext}} \\ \boldsymbol{I}_{\text{supp}}^T\boldsymbol{u} &= 0 \end{aligned} \tag{2}$$

Question 1

To prove that (2) constitutes the necessary and sufficient optimality conditions, we use Theorem 12.1 [1]. Since it is hard to know if LICQ holds in this case, we instead use another possible constraint qualification that is neither weaker nor stronger than LICQ. It says that the condition that all active constraints are linear, is enough.

Because the function $g(\boldsymbol{q})$ and the constraints in (1) are continuously differentiable, the requirements for Thm 12.1 are met.

By using the first of the KKT-conditions,

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0, \tag{3}$$

for both of the variables, namely $\boldsymbol{q}$ and $\boldsymbol{f}^{\text{supp}}$, we get two equations:

$$\begin{aligned} \nabla_{\boldsymbol{q}} \mathcal{L}(\boldsymbol{q}^*, \lambda^*) &= \nabla_{\boldsymbol{q}} g(\boldsymbol{q}^*) - \sum_{i \epsilon \mathcal{A}(\boldsymbol{q}^*)} \lambda_i^* \nabla_{\boldsymbol{q}} c_i(\boldsymbol{q}^*) \\ \nabla_{\boldsymbol{f}^{\text{supp}}} \mathcal{L}(\boldsymbol{f}^{\text{supp}^*}, \lambda^*) &= \nabla_{\boldsymbol{q}} g(\boldsymbol{f}^{\text{supp}^*}) - \sum_{i \epsilon \mathcal{A}(\boldsymbol{f}^{\text{supp}^*})} \lambda_i^* \nabla_{\boldsymbol{f}^{\text{supp}}} c_i(\boldsymbol{f}^{\text{supp}^*}). \end{aligned} \tag{4}$$

This leads to the two equations:

$$\begin{aligned} \boldsymbol{D}\boldsymbol{q} &= \boldsymbol{B}^T\boldsymbol{\lambda} \\ \boldsymbol{I}_{\text{supp}}^T\boldsymbol{\lambda} &= 0 \end{aligned} \tag{5}$$

By choosing λ to be $\boldsymbol{u}$ we see that the KKT-conditions for (1) and the system (2) are equivalent, which is what we aimed to prove. Thm 12.1 is, however, only for proving necessary conditions, but since the KKT-conditions are sufficient under convexity, we can prove it is sufficient as well. For this to be true, $g(\boldsymbol{q})$ must be convex and c_i affine. The first requirement holds since we have a quadratic function. Because the c_i's are linear functions, they are both convex and concave (affine). Hence we have proved that system (2) constitutes the necessary and sufficient optimality conditions for (1).

Question 2

To show that the problem (1) admits at least one solution, we use Theorem 2 in [2].

Since Ω is non-empty, we must also show that it is closed. From [4] we know that if the function is continuous, the hyperplane is closed. As we have already stated, our function $g(\boldsymbol{q})$ is continuous. Hence Ω is both non-empty and closed.

Proving that $g(\boldsymbol{q})$ is coercive, can be shown in this way,

$$\lim_{\|\boldsymbol{q}\|\to\infty} \frac{1}{2}\sum_{j=1}^{m} \frac{\ell_j q_j^2}{E_j A_j} > \lim_{\|\boldsymbol{q}\|\to\infty} \max_{1\leq j\leq m} \frac{\ell_j \|q_j\|^2}{E_j A_j} \to \infty. \tag{6}$$

The requirement that $g(\boldsymbol{q})$ must be a lower semi-continuous function holds as all continuous functions are both lower- and upper semi-continuous. Hence from Theorem 2, we know that (1) admits at least one global minimum.

To conclude that (2) admits at least one solution under these assumptions we use the fact that (1) has a solution. Since we know that (2) is the optimal solution to (1), we then know that (2) has at least one solution.

Question 3

For proving that the problem (1) admits exactly one solution, we use Proposition 3 in [3]. Ω is both non-empty and closed, as previously shown. As the function in Proposition 3 is quadratic, the proposition will work for our case. As the requirements are fulfilled, the proposition tells us that there is at least one globally optimal solution to (1), but the solution is unique as well, since Ω is convex.

By setting $\boldsymbol{f}^{\text{ext}} = \boldsymbol{0}$, we can find this solution. Looking at (1), we see that $g(\boldsymbol{q})$ is minimized if $\boldsymbol{q}$ is equal to $\boldsymbol{0}$. Hence it is easy to see that this implies that $\boldsymbol{f}^{\text{supp}} = \boldsymbol{0}$ as well, from the constraints. This gives $(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = (\boldsymbol{0}, \boldsymbol{0})$ which is our globally optimal solution.

Equations for questions 4-5

$$
\begin{aligned}
\underset{(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}})}{\text{minimize}} \quad & \frac{1}{2}\sum_{j=1}^{m}\frac{\ell_j q_j^2}{E_j A_j} = g(\boldsymbol{A},\boldsymbol{q}), \\
\text{subject to} \quad & \boldsymbol{B}\boldsymbol{q} = \boldsymbol{I}_{\text{supp}}\boldsymbol{f}^{\text{supp}} + \boldsymbol{I}_{\text{ext}}\boldsymbol{f}^{\text{ext}}, \\
& \sum_{j=1}^{m}\rho_j \ell_j Aj \leq M, \\
& \underline{A}_j \leq A_j \leq \overline{A}_j, \qquad j = 1, ..., m.
\end{aligned} \tag{7}
$$

Question 4

It is possible to use Theorem 2 [2] for system (7) as well. Whenever feasible means that Ω is non-empty and since the constraints have no strict inequalities, we know Ω is closed. What remains to show is that $g(\boldsymbol{A},\boldsymbol{q})$ is coercive and lower semi-continuous.

First we check if it is coercive. As $g(\boldsymbol{A},\boldsymbol{q})$ is a function of two variables, we must check if it is coercive when both $\|\boldsymbol{A}\|$ and $\|\boldsymbol{q}\|$ goes to infinity. Since $\boldsymbol{q}$ is dependent on $\boldsymbol{f}^{\text{supp}}$, and $\|\boldsymbol{f}^{\text{supp}}\|$ goes to infinity, we know it's no problem that $\|\boldsymbol{q}\|$ does it as well. We divide it in two cases, when the element in $\boldsymbol{q}$ and $\boldsymbol{A}$ that goes to infinity, namely q_k and A_p are equal indices, or different indices. We will start with the case where k = p. Notice that in the last limit, $\|\boldsymbol{q}\|_\infty = q_k$, $\|\boldsymbol{A}\|_\infty = q_p$.

$$
\text{k = p}: \lim_{\substack{\|\boldsymbol{q}\|\to\infty \\ \|\boldsymbol{A}\|\to\infty \\ \|\boldsymbol{f}^{\text{supp}}\|\to\infty}} \frac{1}{2}\sum_{j=1}^{m}\frac{\ell_j q_j^2}{E_j A_j} = \frac{1}{2}\sum_{\substack{j=1 \\ j\neq k}}^{m}\frac{\ell_j q_j^2}{E_j A_j} + \lim_{\substack{q_k\to\infty \\ A_k\to\infty \\ f^{\text{supp}_k}\to\infty}} \frac{\ell_k {q_k}^2}{E_k A_k} = \infty. \tag{8}
$$

$$
\text{k} \neq \text{p}: \lim_{\substack{\|\boldsymbol{q}\|\to\infty \\ \|\boldsymbol{A}\|\to\infty \\ \|\boldsymbol{f}^{\text{supp}}\|\to\infty}} \frac{1}{2}\sum_{j=1}^{m}\frac{\ell_j q_j^2}{E_j A_j} = \frac{1}{2}\sum_{\substack{j=1 \\ j\neq k \\ j\neq p}}\frac{\ell_j q_j^2}{E_j A_j} + \lim_{\substack{q_k\to\infty \\ \|\boldsymbol{f}^{\text{supp}}\|\to\infty}} \frac{\ell_k {q_k}^2}{E_k A_k} + \lim_{A_p\to\infty}\frac{\ell_p {q_p}^2}{E_p A_p} = \infty. \tag{9}
$$

Lastly, $g(\boldsymbol{A},\boldsymbol{q})$ is lower semi-continuous because it is continuous. Hence theorem 2 applies, and we have at least one global minimum.

We now check to see if the theorem still holds when $\underline{A}_j = 0$. The constraint now becomes $A_j \geq 0$. If $q_j \neq 0$ and $\frac{q_j^2}{0} = +\infty$ we see that $g(0, q)$ is equal to infinity. If, on the other hand, $q_j = 0$, we get $\frac{0}{0}$ inside the sum, which is equal to 0. Hence we get a value $\rho \neq 0$. We see that the first case is still continuous, while the second is lower semi-continuous, as it drops from ∞ when q_j goes to 0, while it is ρ when it's equal to 0. Thus theorem 2 still applies.

Question 5

The KKT-conditions for (7) are as follows:

$$
\begin{aligned}
\boldsymbol{D}\boldsymbol{q} &= \boldsymbol{B}^T\boldsymbol{\lambda} \\
\boldsymbol{I}_{\text{supp}}^T\boldsymbol{\lambda} &= 0 \\
\boldsymbol{K} - \boldsymbol{P}\boldsymbol{\mu} - \boldsymbol{I}\boldsymbol{\gamma} + \boldsymbol{I}\boldsymbol{\xi} &= 0 \\
\mu_j &\geq 0 \text{ for } j = 1, ..., m \\
\gamma_j &\geq 0 \text{ for } j = 1, ..., m \\
\xi_j &\geq 0 \text{ for } j = 1, ..., m \\
\mu_j(\rho_j l_j A_j - M) &= 0 \text{ for } j = 1, ..., m \\
\gamma_j(A_j - \overline{A_j}) &= 0 \text{ for } j = 1, ..., m \\
\xi_j(\underline{A_j} - A_j) &= 0 \text{ for } j = 1, ..., m
\end{aligned}
\tag{10}
$$

$$
\begin{aligned}
\boldsymbol{K} &= \left[-\tfrac{l_1 q_1^2}{2E_1 A_1^2}, -\tfrac{l_2 q_2^2}{2E_2 A_2}, \dots, -\tfrac{l_m q_m^2}{2E_m A_m}\right]^T \\
\boldsymbol{P} &= \begin{bmatrix} \rho_1 l_1 & 0 & \dots & 0 \\ 0 & \rho_2 l_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \rho_m l_m \end{bmatrix} \\
\boldsymbol{I} &= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}
\end{aligned}
$$

$$
\bigtriangledown^2 g(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{supp}) = \left[\begin{array}{cc|cc|cc}
\frac{l_1 q_1^2}{E_1 A_1^3} & 0 & -\frac{l_1 q_1}{E_1 A_1^2} & 0 & 0 & 0 \\
0 & \ddots & 0 & \ddots & 0 & \ddots \\
\hline
-\frac{l_1 q_1}{E_1 A_1^2} & 0 & \frac{l_1}{E_1 A_1} & 0 & 0 & 0 \\
0 & \ddots & 0 & \ddots & 0 & \ddots \\
\hline
0 & 0 & 0 & 0 & 0 & 0 \\
0 & \ddots & 0 & \ddots & 0 & \ddots
\end{array}\right]
$$

If we let $\boldsymbol{w} \epsilon \Omega$, we can observe that

$$\boldsymbol{w}^T \bigtriangledown^2 g(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{supp})\boldsymbol{w} =$$

$$\begin{bmatrix} A_1 \dots A_m & q_1 \dots q_m \end{bmatrix} \bigtriangledown^2 g(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{supp}) \begin{bmatrix} A_1 \\ \vdots \\ \vdots \\ A_m \\ q_1 \\ \vdots \\ \vdots \\ q_m \end{bmatrix} = 0,$$

and thus $g(\boldsymbol{A}, \boldsymbol{q})$ is positive semi-definite, which is a sufficient condition for convexity. Since all our optimality constraints are affine, then according to [5], (10) is both necessary and sufficient for optimality for (7).

Equations for questions 6-7

$$\begin{aligned} \underset{(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}})}{\text{minimize}} \quad & \frac{1}{2}\sum_{j=1}^{m} \frac{\ell_j q_j^2}{E_j A_j} - \mu \log[M - \sum_{j=1}^{m} \rho_j l_j A_j] \\ & -\mu \sum_{j=1}^{m} \log[(A_j - \underline{A_j})(\overline{A_j} - A_j)] = g(\boldsymbol{A}, \boldsymbol{q}), \\ \text{subject to} \quad & \boldsymbol{B}\boldsymbol{q} = \boldsymbol{I}_{\text{supp}}\boldsymbol{f}^{\text{supp}} + \boldsymbol{I}_{\text{ext}}\boldsymbol{f}^{\text{ext}}, \end{aligned} \tag{11}$$

Question 6

The KKT-conditions for (7) are as follows:

$$\begin{aligned} \boldsymbol{D}\boldsymbol{q} &= \boldsymbol{B}^T\boldsymbol{\lambda} \\ \boldsymbol{I}_{\text{supp}}^T\boldsymbol{\lambda} &= 0 \\ \boldsymbol{K} - \boldsymbol{P}\boldsymbol{\mu} - \boldsymbol{I}\boldsymbol{\gamma} + \boldsymbol{I}\boldsymbol{\xi} &= 0 \\ \mu_j &\geq 0 \text{ for } j = 1, ..., m \\ \gamma_j &\geq 0 \text{ for } j = 1, ..., m \\ \xi_j &\geq 0 \text{ for } j = 1, ..., m \\ \mu_j(\rho_j l_j A_j - M) &= 0 \text{ for } j = 1, ..., m \\ \gamma_j(A_j - \overline{A_j}) &= 0 \text{ for } j = 1, ..., m \\ \xi_j(\underline{A_j} - A_j) &= 0 \text{ for } j = 1, ..., m \end{aligned} \tag{12}$$

References

[1] Jorge Nocedal, Stephen J. Wright, *Numerical Optimization*, 2nd Edition, 2006.

[2] Anton Evgrafov, *Optimization Theory*

[3] Anton Evgrafov, *Introduction to optimality conditions*

[4] California Institute of Technology, *Separating Hyperplane Theorems* http://www.hss.caltech.edu/~kcb/Notes/SeparatingHyperplane.pdf, 28. March 2014

[5] Anton Evgrafov, *Theorem: KKT conditions are sufficient under convexity*