SEMESTER PROJECT

Candidates 10006 and 10030

Department of Mathematical Sciences
Norwegian University of Science and Technology

April 4, 2014

We refer to [4] for a full description of the project, questions and mentioned equations/problems.

Question 1

Let $\varphi(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}})$ denote the objective function and $\boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}})$ express the equality constraints, that is,

$$\varphi(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = \frac{1}{2}\sum_{j=1}^{m} \frac{\ell_j q_j^2}{E_j A_j} = \frac{1}{2}\boldsymbol{q}^\top \boldsymbol{D}\boldsymbol{q} \quad \text{and} \quad \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = \boldsymbol{B}\boldsymbol{q} - \boldsymbol{I}_{\text{supp}}\boldsymbol{f}^{\text{supp}} - \boldsymbol{I}_{\text{ext}}\boldsymbol{f}^{\text{ext}}.$$

We form the Lagrangian

$$\mathcal{L}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) = \varphi(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) - \boldsymbol{u}^\top \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}),$$

where $\boldsymbol{u} \in \mathbb{R}^{3n}$ denotes the vector of Lagrange multipliers, and observe readily that

$$\nabla_{\boldsymbol{q}}\mathcal{L}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) = \boldsymbol{D}\boldsymbol{q} - \boldsymbol{B}^\top\boldsymbol{u} \quad \text{and} \quad \nabla_{\boldsymbol{f}^{\text{supp}}}\mathcal{L}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) = \boldsymbol{I}_{\text{supp}}^\top \boldsymbol{u}.$$

The Karush-Kuhn-Tucker (KKT) conditions for the minimization problem (5) are then given as

$$\begin{aligned} \nabla_{\boldsymbol{q}}\mathcal{L}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) &= \boldsymbol{0}, \\ \nabla_{\boldsymbol{f}^{\text{supp}}}\mathcal{L}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) &= \boldsymbol{0}, \\ \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) &= \boldsymbol{0}, \end{aligned} \qquad \text{which is the same as} \qquad \begin{aligned} \boldsymbol{D}\boldsymbol{q} &= \boldsymbol{B}^\top\boldsymbol{u}, \\ \boldsymbol{I}_{\text{supp}}^\top\boldsymbol{u} &= \boldsymbol{0}, \\ \boldsymbol{B}\boldsymbol{q} &= \boldsymbol{I}_{\text{supp}}\boldsymbol{f}^{\text{supp}} + \boldsymbol{I}_{\text{ext}}\boldsymbol{f}^{\text{ext}}. \end{aligned}$$

Note that the complementarity conditions are automatically satisfied, and hence omitted, since there are no inequality constraints.

Observe that both φ and $\boldsymbol{c}$ are continuously differentiable. Moreover, $\boldsymbol{c}$ expresses affine equality constraints, and so Proposition 6 in [3] implies that the KKT conditions are necessary optimality conditions for the minimization problem (5).

In addition, the objective function is convex, which can be seen by rewriting φ as

$$\varphi(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = \frac{1}{2}[\boldsymbol{q}^\top \; \boldsymbol{f}^{\text{supp}\top}] \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{q} \\ \boldsymbol{f}^{\text{supp}} \end{bmatrix}$$

and noticing that $\boldsymbol{D}$ is a symmetric positive definite matrix. As the constraints are affine functions, the feasible set is convex, and hence "The Convex KKT Theorem" in [5] gives the sufficiency of the KKT conditions for problem (5).

Question 2

As φ is continuously differentiable, it is also *lower semi-continuous*. Furthermore, to prove the *coercivity* of φ, note that letting $\|\boldsymbol{f}^{\text{supp}}\| \to \infty$ forces $\|\boldsymbol{q}\| \to \infty$ in light of the equality constraint. Hence coercivity follows from

$$\varphi(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = \frac{1}{2}\sum_{j=1}^{m} \frac{\ell_j q_j^2}{E_j A_j} \geq \frac{1}{2} \min_{j=1}^{m} \frac{\ell_j}{E_j A_j} \|\boldsymbol{q}\|^2 \to \infty \quad \text{as} \quad \|(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}})\| \to \infty.$$

By assumption the feasible set $\Omega = \{(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) \in \mathbb{R}^m \times \mathbb{R}^{3n_s} : \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = \boldsymbol{0}\}$ is nonempty, and since $\boldsymbol{c}$ is an affine function, Ω is also *closed*.

With these properties at hand, Theorem 2 in [2] then implies the existence of at least one solution to the minimization problem (5). Existence of a solution of the KKT conditions (4) then follows from the *necessity* of the optimality conditions.

Question 3

Assume that $(\boldsymbol{q}_1, \boldsymbol{f}_1^{\text{supp}})$ and $(\boldsymbol{q}_2, \boldsymbol{f}_2^{\text{supp}})$ are two solutions to the minimization problem (5). By the equivalence between the minimization problem (5) and the KKT conditions (4), we obtain via simple subtraction that

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{q}_1 - \boldsymbol{q}_2) &= \boldsymbol{B}^\top(\boldsymbol{u}_1 - \boldsymbol{u}_2), \\ \boldsymbol{I}_{\text{supp}}^\top(\boldsymbol{u}_1 - \boldsymbol{u}_2) &= \boldsymbol{0}, \\ \boldsymbol{B}(\boldsymbol{q}_1 - \boldsymbol{q}_2) &= \boldsymbol{I}_{\text{supp}}(\boldsymbol{f}_1^{\text{supp}} - \boldsymbol{f}_2^{\text{supp}}), \end{aligned}$$

where $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are the corresponding Lagrange multipliers associated with $(\boldsymbol{q}_1, \boldsymbol{f}_1^{\text{supp}})$ and $(\boldsymbol{q}_2, \boldsymbol{f}_2^{\text{supp}})$, respectively. This gives that

$$\begin{aligned} (\boldsymbol{q}_1 - \boldsymbol{q}_2)^\top \boldsymbol{D}(\boldsymbol{q}_1 - \boldsymbol{q}_2) &= (\boldsymbol{q}_1 - \boldsymbol{q}_2)^\top \boldsymbol{B}^\top(\boldsymbol{u}_1 - \boldsymbol{u}_2) \\ &= [\boldsymbol{B}(\boldsymbol{q}_1 - \boldsymbol{q}_2)]^\top(\boldsymbol{u}_1 - \boldsymbol{u}_2) \\ &= [\boldsymbol{I}_{\text{supp}}(\boldsymbol{f}_1^{\text{supp}} - \boldsymbol{f}_2^{\text{supp}})]^\top(\boldsymbol{u}_1 - \boldsymbol{u}_2) \\ &= (\boldsymbol{f}_1^{\text{supp}} - \boldsymbol{f}_2^{\text{supp}})^\top \boldsymbol{I}_{\text{supp}}^\top(\boldsymbol{u}_1 - \boldsymbol{u}_2) \\ &= \boldsymbol{0}, \end{aligned}$$

and so $\boldsymbol{q}_1 - \boldsymbol{q}_2 = \boldsymbol{0}$ since $\boldsymbol{D}$ is a symmetric positive definite matrix and hence induces a norm $\|\boldsymbol{q}\|_D^2 := \boldsymbol{q}^\top D \boldsymbol{q}$. Thus $\boldsymbol{q}_1 = \boldsymbol{q}_2$, and so $\boldsymbol{I}_{\text{supp}}(\boldsymbol{f}_1^{\text{supp}} - \boldsymbol{f}_2^{\text{supp}}) = \boldsymbol{0}$ as well. But this implies that $\boldsymbol{f}_1^{\text{supp}} - \boldsymbol{f}_2^{\text{supp}} = \boldsymbol{0}$ by the construction of $\boldsymbol{I}_{\text{supp}}$. Uniqueness of the minimization problem (5) therefore follows.

Question 4

Since the objective function $\varphi(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}})$ (now considered also a function of $\boldsymbol{A}$) is continuously differentiable, it is also *lower semi-continuous*. The feasible set may be written as $\Omega = \Omega_1 \cap \Omega_2 \cap \Omega_3$, with

$$\begin{aligned}
\Omega_1 &= \{(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) \in (\mathbb{R}\cup\{\infty\})^m \times \mathbb{R}^m \times \mathbb{R}^{3n_s} : \boldsymbol{c}(\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) = \boldsymbol{0}\},\\
\Omega_2 &= \{(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) \in (\mathbb{R}\cup\{\infty\})^m \times \mathbb{R}^m \times \mathbb{R}^{3n_s} : c_0(\boldsymbol{A}) \ge 0\}\\
\text{and } \Omega_3 &= \{(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) \in (\mathbb{R}\cup\{\infty\})^m \times \mathbb{R}^m \times \mathbb{R}^{3n_s} : c_j(\boldsymbol{A}) \ge 0, j = 1,\dots,2m\},
\end{aligned}$$

where

$$c_0(\boldsymbol{A}) = M - \sum_{j=1}^m \rho_j \ell_j A_j \quad \text{and} \quad c_j(\boldsymbol{A}) = \begin{cases} A_j - \underline{A}_j & \text{if } j = 1,\dots,m;\\ \overline{A}_{j-m} - A_{j-m} & \text{if } j = m+1,\dots,2m. \end{cases}$$

The set $\{(\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) \in \mathbb{R}^m \times \mathbb{R}^{3n_s} : \boldsymbol{c}(\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) = \boldsymbol{0}\}$ was shown to be closed in Question 2, and so by extension it follows readily that Ω_1 is closed. Moreover, Ω_2 defines a closed half-space, while Ω_3 is geometrically a closed box. Since the intersection of closed sets is closed, it follows that Ω is *closed*.

To prove *coercivity* of φ, note that $\boldsymbol{A}$ is bounded below from Ω_3 and bounded above via Ω_2. Hence letting $\|(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}})\| \to \infty$ forces $\|(\boldsymbol{q},\boldsymbol{f}^{\text{supp}})\| \to \infty$, and from the equality constraints Ω_1 as in Question 2, this again implies $\|\boldsymbol{q}\| \to \infty$. Thus

$$\varphi(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) = \frac{1}{2}\sum_{j=1}^m \frac{\ell_j q_j^2}{E_j A_j} \ge \frac{1}{2}\min_{j=1}^m \frac{\ell_j}{E_j A_j}\|\boldsymbol{q}\|^2 \to \infty \quad \text{as} \quad \|(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}})\| \to \infty.$$

By assumption the feasible set is nonempty, and so with all these properties combined, Theorem 2 in [2] implies the existence of at least one optimal solution to problem (6).

We now consider that possibility that $\underline{A}_j = 0$, with the understanding that $q_j^2/0 = \infty$, if $q_j \ne 0$ and 0, if $q_j = 0$. This enlarges the sets Ω_2 and Ω_3, but they will again be a closed half-space and a closed box, respectively. Hence Ω remains closed. Since $\|(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}})\| \to \infty$ implies $\|\boldsymbol{q}\| \to \infty$, the coercivity of φ still holds.

To check the lower semi-continuity of φ, we first analyze the simpler function $h(A_j, q_j) = q_j^2/A_j$. On the part of the domain where $A_j > 0$, it is quickly seen that h is continuously differentiable, and thus also lower semi-continuous. When $A_j = 0$, the analysis has to be more delicate. Let $\{(A_j^n, q_j^n)\}_n$ be a sequence converging to $(0, q_j)$, with $q_j \ne 0$. Since $q_j \ne 0$, eventually we must have $q_j^n \ne 0$ for large n. Hence, as $A_j^n \to 0$, we get that

$$\liminf_{n\to\infty} h(A_j^n, q_j^n) = \infty = h(0, q_j).$$

For the situation with $q_j = 0$, we distinguish between two cases. If eventually $q_j^n = 0$ for all large n, then $\lim_n h(A_j^n, q_j^n) = 0 = h(0,0)$, and we are done. For the other case, assume that $q_j^n \ne 0$ for infinitely many n. Then readily we have $\liminf_n h(A_j^n, q_j^n) \ge 0 = h(0,0)$ by the mere fact that $h \ge 0$ on the domain.

Thus in all situations $\liminf_n h(A_j^n, q_j^n) \ge h(A_j, q_j)$, and so the function h is lower semi-continuous by Proposition 1 in [2]. Moreover, because a conical combination (nonnegative weighted sum) of lower semi-continuous functions is lower semi-continuous via Proposition 2.13 a) in [1], it follows that φ is lower semi-continuous as a function of $(\boldsymbol{A}, \boldsymbol{q})$. Since φ does not explicitly depend on $\boldsymbol{f}^{\text{supp}}$, we conclude that φ is lower semi-continuous. In total the above existence result still holds for the optimal solution to problem (6).

Question 5

We form the Lagrangian

$$\mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) = \varphi(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) - \tilde{\boldsymbol{u}}^\top \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) - \sum_{j=0}^{2m} u_{3n+1+j} c_j(\boldsymbol{A}),$$

where $\boldsymbol{u} = (u_j) \in \mathbb{R}^{3n+1+2m}$ denotes the vector of Lagrange multipliers, and $\tilde{\boldsymbol{u}} \in \mathbb{R}^{3n}$ constitutes the first $3n$ components of $\boldsymbol{u}$. The KKT optimality conditions for problem (6) then reads

$$\begin{aligned} \nabla_{\boldsymbol{A}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) &= \boldsymbol{0}, \\ \nabla_{\boldsymbol{q}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) &= \boldsymbol{0}, \\ \nabla_{\boldsymbol{f}^{\text{supp}}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) &= \boldsymbol{0}, \end{aligned} \qquad \begin{aligned} \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) &= \boldsymbol{0}, \\ c_j(\boldsymbol{A}) &\geq 0, j = 0, \ldots, 2m, \\ u_{3n+1+j} &\geq 0, j = 0, \ldots, 2m, \\ u_{3n+1+j} c_j(\boldsymbol{A}) &= 0, j = 0, \ldots, 2m, \end{aligned}$$

which is the same as

$$\begin{aligned} \nabla_{\boldsymbol{A}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}) &= \boldsymbol{0}, \\ \boldsymbol{D}\boldsymbol{q} &= \boldsymbol{B}^\top \tilde{\boldsymbol{u}}, \\ \boldsymbol{I}_{\text{supp}}^\top \tilde{\boldsymbol{u}} &= \boldsymbol{0}, \end{aligned} \qquad \begin{aligned} & \boldsymbol{B}\boldsymbol{q} = \boldsymbol{I}_{\text{supp}} \boldsymbol{f}^{\text{supp}} + \boldsymbol{I}_{\text{ext}} \boldsymbol{f}^{\text{ext}}, \\ & \sum_{j=1}^{m} \rho_j \ell_j A_j \leq M, \\ & \underline{A}_j \leq A_j \leq \overline{A}_j, j = 1, \ldots, m, \\ & u_{3n+1+j} \geq 0, j = 0, \ldots, 2m, \\ & u_{3n+1+j} c_j(\boldsymbol{A}) = 0, j = 0, \ldots, 2m, \end{aligned}$$

where some of the constraints have been written compactly and

$$(\nabla_{\boldsymbol{A}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}))_j = -\frac{1}{2} \frac{\ell_j q_j^2}{E_j A_j^2} + u_{3n+1} \rho_j \ell_j - u_{3n+1+j} + u_{3n+1+m+j}, \quad j = 1, \ldots, m.$$

The objective function and all the constraints are readily seen to be continuously differentiable. Since all the constraints are affine functions, Lemma 12.7 in [6] implies that the linearized feasible direction set at a local optimal solution equals the related tangent cone, and hence provides a sufficient constraint qualification. By Theorem 12.1 in [6] the KKT optimality conditions are then necessary for the optimality of problem (6).

Moreover, as the constraints are affine, the feasible set is convex. Combining these two properties, the sufficiency of the KKT conditions follows from "The Convex KKT Theorem" in [5] if we can prove that the objective function is convex. To this end, note that the Hessian of the function $h(q_j, A_j) = q_j^2/A_j$ equals

$$\begin{bmatrix} \dfrac{2}{A_j} & -\dfrac{2q_j}{A_j^2} \\ -\dfrac{2q_j}{A_j^2} & \dfrac{2q_j^2}{A_j^3} \end{bmatrix} \qquad \text{with eigenvalues 0 and } \frac{2}{A_j}\left(1+\frac{q_j^2}{A_j^2}\right),$$

and so h is convex. Since a conical combination (nonnegative weighted sum) of convex functions is convex, it follows that φ is convex as a function of $(\boldsymbol{A}, \boldsymbol{q})$. Because $\boldsymbol{f}^{\text{supp}}$ does not occur explicitly in φ, the objective function $\varphi(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}})$ is convex on the feasible set.

Question 6

The Lagrangian reads $\mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \hat{\boldsymbol{u}}) = \psi(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) - \hat{\boldsymbol{u}}^\top \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}})$, where the objective function ψ is defined as

$$\psi(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = \varphi(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) - \mu \log\left[M - \sum_{j=1}^{m} \rho_j \ell_j A_j\right] - \mu \sum_{j=1}^{m} \log[(A_j - \underline{A}_j)(\overline{A}_j - A_j)],$$

and $\hat{\boldsymbol{u}}$ is the vector of Lagrange multipliers. The KKT optimality conditions for problem (7) are

$$\begin{aligned} \nabla_{\boldsymbol{A}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \hat{\boldsymbol{u}}) &= \boldsymbol{0}, \\ \nabla_{\boldsymbol{q}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \hat{\boldsymbol{u}}) &= \boldsymbol{0}, \\ \nabla_{\boldsymbol{f}^{\text{supp}}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \hat{\boldsymbol{u}}) &= \boldsymbol{0}, \\ \boldsymbol{c}(\boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) &= \boldsymbol{0}, \end{aligned} \qquad \text{which is the same as} \qquad \begin{aligned} &\nabla_{\boldsymbol{A}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \hat{\boldsymbol{u}}) = \boldsymbol{0}, \\ &\boldsymbol{D}\boldsymbol{q} = \boldsymbol{B}^\top \hat{\boldsymbol{u}}, \\ &\boldsymbol{I}_{\text{supp}}^\top \hat{\boldsymbol{u}} = \boldsymbol{0}, \\ &\boldsymbol{B}\boldsymbol{q} = \boldsymbol{I}_{\text{supp}} \boldsymbol{f}^{\text{supp}} + \boldsymbol{I}_{\text{ext}} \boldsymbol{f}^{\text{ext}}, \end{aligned}$$

where

$$(\nabla_{\boldsymbol{A}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \hat{\boldsymbol{u}}))_j = -\frac{1}{2}\frac{\ell_j q_j^2}{E_j A_j^2} + \frac{\mu \rho_j \ell_j}{M - \sum_{i=1}^{m} \rho_i \ell_i A_i} - \frac{\mu}{A_j - \underline{A}_j} + \frac{\mu}{\overline{A}_j - A_j}, \quad j = 1, \ldots, m.$$

A comparison with the optimality conditions for the problem (6) shows that equality constraint condition is exactly the same. Moreover, the equations coming from the gradient of the Lagrangian with respect to $\boldsymbol{q}$ and $\boldsymbol{f}^{\text{supp}}$ are almost identical for the problems (6) and (7). And we note that the inequality constraints and their corresponding Lagrange multipliers from problem (6) have moved into the equation involving the gradient of the Lagrangian with respect to $\boldsymbol{A}$ for problem (7). Since

$$(\nabla_{\boldsymbol{A}} \mathcal{L}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}, \boldsymbol{u}))_j = -\frac{1}{2}\frac{\ell_j q_j^2}{E_j A_j^2} + u_{3n+1}\rho_j \ell_j - u_{3n+1+j} + u_{3n+1+m+j}, \quad j = 1, \ldots, m.$$

in problem (6) and all the Lagrange multipliers corresponding to inequality constraints are nonnegative, reasonable approximations for these Lagrange multipliers in terms of an optimal solution to (7) are

$$u_{3n+1} \approx \frac{\mu}{M - \sum_{i=1}^{m} \rho_i \ell_i A_i}, \quad u_{3n+1+j} \approx \frac{\mu}{A_j - \underline{A}_j}, \quad u_{3n+1+m+j} \approx \frac{\mu}{\overline{A}_j - A_j}, \quad j = 1, \ldots, m.$$

The equality constraints function $\boldsymbol{c}$ is as before continuously differentiable, and for practical purposes the same holds also for the objective function ψ. We note that there may be issues about differentiability of ψ when some of the logarithmic parts assume the value ∞, that is, some of the inequality constraints from Question 5 are active. In fact, the objective function may spit out complex values. However, when it comes to minimizing ψ, this causes no problems since $(-\log\gamma) \to \infty$ as $\gamma \searrow 0$. Because $\boldsymbol{c}$ expresses affine functions, Lemma 12.7 in [6] implies that the linearized feasible direction set at a local optimal solution equals the related tangent cone, and hence provides a sufficient constraint qualification. By Theorem 12.1 in [6] the KKT optimality conditions are then necessary for the optimality of problem (7).

From Question 5 we know that φ is convex, and $(-\log(\cdot))$ is a convex function since the its second derivative is positive. Now, the arguments in the logarithmic parts of the objective funtion are all affine (split the last sum in two by the product rule for logarithms), and hence also convex. Since the composition of an affine function with a convex function is convex[1], it follows that all the negative logarithms are convex. Thus, as the conical combination of convex functions is convex, we obtain that ψ is convex.

The minimization problem (7) is thus convex, with affine equality constraints, and so "The Convex KKT Theorem" in [5] implies that the KKT conditions are sufficient optimality conditions for problem (7).

Question 7

To implement a linesearch SQP algorithm for the barrier problem (7), we use Algorithm 18.3 outlined in [6] based on the exact Hessian of the Lagrangian and with ℓ_1 merit function and backtracking (Armijo) linesearch.

The gradient of the objective function reads

$$\nabla\psi(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) = \begin{bmatrix} \nabla_{\boldsymbol{A}}\psi(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) \\ \nabla_{\boldsymbol{q}}\psi(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) \\ \nabla_{\boldsymbol{f}^{\text{supp}}}\psi(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}}) \end{bmatrix} = \begin{bmatrix} \nabla_{\boldsymbol{A}}\mathcal{L}(\boldsymbol{A},\boldsymbol{q},\boldsymbol{f}^{\text{supp}},\hat{\boldsymbol{u}}) \\ \boldsymbol{D}\boldsymbol{q} \\ \boldsymbol{0} \end{bmatrix},$$

and we calculate the exact Hessian of the Lagrangian to be a $(2m+3n_s)\times(2m+3n_s)$ matrix which takes the form

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_{\boldsymbol{A},\boldsymbol{A}} & \mathcal{H}_{\boldsymbol{q},\boldsymbol{A}} & \mathcal{H}_{\boldsymbol{f}^{\text{supp}},\boldsymbol{A}} \\ \mathcal{H}_{\boldsymbol{A},\boldsymbol{q}} & \mathcal{H}_{\boldsymbol{q},\boldsymbol{q}} & \mathcal{H}_{\boldsymbol{f}^{\text{supp}},\boldsymbol{q}} \\ \mathcal{H}_{\boldsymbol{A},\boldsymbol{f}^{\text{supp}}} & \mathcal{H}_{\boldsymbol{q},\boldsymbol{f}^{\text{supp}}} & \mathcal{H}_{\boldsymbol{f}^{\text{supp}},\boldsymbol{f}^{\text{supp}}} \end{bmatrix},$$

where

$$\mathcal{H}_{\boldsymbol{A},\boldsymbol{A}} = \left(\frac{\mu\rho_i\ell_i\rho_j\ell_j}{(M-\sum_{k=1}^m \rho_k\ell_k A_k)^2}\right)_{ij} + \operatorname{diag}\left(\frac{\ell_j q_j^2}{E_j A_j^3} + \frac{\mu}{(A_j-\underline{A}_j)^2} + \frac{\mu}{(\overline{A}_j - A_j)^2}, j=1,\dots,m\right),$$

$$\mathcal{H}_{\boldsymbol{A},\boldsymbol{q}} = \mathcal{H}_{\boldsymbol{q},\boldsymbol{A}} = \operatorname{diag}\left(-\frac{\ell_j q_j}{E_j A_j^2}, j=1,\dots,m\right) \quad \text{and} \quad \mathcal{H}_{\boldsymbol{q},\boldsymbol{q}} = \boldsymbol{D}$$

[1]We observe that $(f\circ g)(\theta x + (1-\theta)y) = f(\theta g(x) + (1-\theta)g(y)) \le \theta(f\circ g)(x) + (1-\theta)(f\circ g)(y)$ for all x, y in the domain of g and $\theta \in (0,1)$ if f is convex and g is affine.

are the only nonzero matrix blocks in $\mathcal{H}$. Moreover, the constraints Jacobian equals

$$\nabla \boldsymbol{c}(\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}}) = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{B} & -\boldsymbol{I}_{\text{supp}} \end{bmatrix}.$$

We have tested numerically that Assumptions 18.1 in [6] hold for some of the test cases below, but are not sure if the constraints Jacobian has full row rank in general.

The programming environment MATLAB has been used, and we have stribed to make the algorithm independent of the specific problem. For example, we introduce a vector $\boldsymbol{x} = (\boldsymbol{A}, \boldsymbol{q}, \boldsymbol{f}^{\text{supp}})$, and invoke function handles for the objective function, the equality constraints, the objective gradient and the Hessian of the Lagrangian into the SQP algorithm. Moreover, systematic use of global variables in the problem specific functions enables effective implementations. In particular, for calculation of the Hessian of the Lagrangian, we have exploited a fast outer product structure when building $\mathcal{H}_{\boldsymbol{A},\boldsymbol{A}}$:

```
tmp = rho .* l;
tmp = tmp ./ (M - dot(tmp, x(1:m)));

H_AA =    mu * (tmp * tmp') ...
      + diag(((l .* x(m+1:2*m).^2) ./ (E .* x(1:m).^3)) ...
            + mu*((x(1:m) - A_underline).^(-2) + (A_overline - x(1:m)).^(-2)));
```

The search directions $\boldsymbol{p}_k$ and $\boldsymbol{p}_\lambda$ are found by solving the Newton-KKT system (18.6) in [6] via the backslash-operator in MATLAB. Moreover, the penalty parameter μ_k is calculated as a modified combination of the inequalities (18.32) and (18.36) in [6]:

```
mu_lower_lim = (dot(f_grad_x, p) + .5 * sigma * dot(p, H_x * p)) / ((1 - ...
    rho) * norm(c_eq_x, 1));
if mu_lower_lim > 0;
  mu = 1.1 * mu_lower_lim;
else
  mu = 1.1 * max(norm(lambda, Inf), norm(lambda + p_lambda, Inf));
end
```

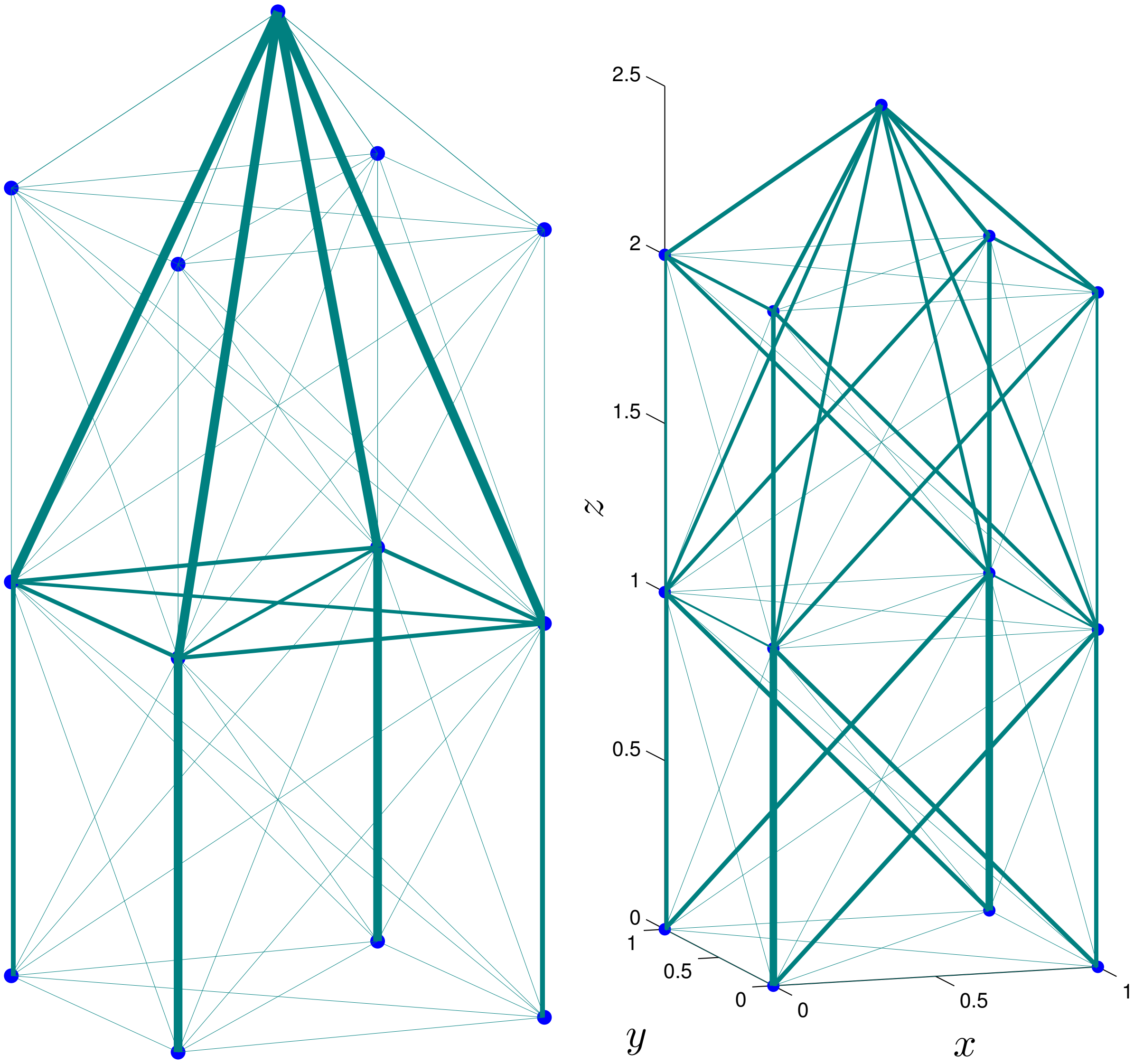
To be more detailed, inequality (18.36) is preferred if its righthand-side is positive, and we let $\mu_k > \max(\|\boldsymbol{\lambda}_k\|_\infty, \|\boldsymbol{\lambda}_k + \boldsymbol{p}_\lambda\|_\infty)$ otherwise. The implementation uses $\sigma = 1$ and $\rho = 0.5$.

When it comes to the backtracking procedure, it is important to guarantee that the input arguments in the ℓ_1 penalty function are valid, in the sense that the arguments in the logarithms coming from the objective function are all strictly positive. Hence we first shorten (repeatedly halving) the step length α_k until the inequality constrants from problem (6) are satisfied with strict inequalities. From then on the ordinary backtracking method takes over.

As a convergence criterion for the SQP algorithm, we use the norm of the righthand-side of the Newton-KKT system (18.6) in [6], which measures the violation of both the KKT conditions and the equality constraints.

To test the SQP algorithm, four test examples are considered. In the construction of these, we make use of the MATLAB-functions `find`, `intersect` and `setdiff` to efficiently locate specific nodes in the truss. We mention that in the first three examples, all the problem settings as listed in [4] are used.

Examples 1: Tower, and 2: Cantilever



(a) Truss example: Tower. **(b)** Truss example: Cantilever.

Figure 1: Numerical solutions of truss example 1 and 2.

With the given problem settings, we find that the values for the starting points have to be chosen quite carefully. We control-check that the inequality constraints from problem (6) are all satisfied before entering the SQP algorithm. Especially $\boldsymbol{A}$ has to be quite close to to $\underline{\boldsymbol{A}}$ componentvise for the initial point. Details can be seen in the attached MATLAB-files. We have managed identified the solution with an error as low as $\sim 10^{-13}$ in 25 iterations for Example 1, and the solution to both Example 1 and Example 2 can be seen in Figure 1.

With a roughly best as possible numerical solution from Example 1, we have estimated the order of convergence. As can be seen in Figure 2, we seem to obtain a qubic rate of convergence.

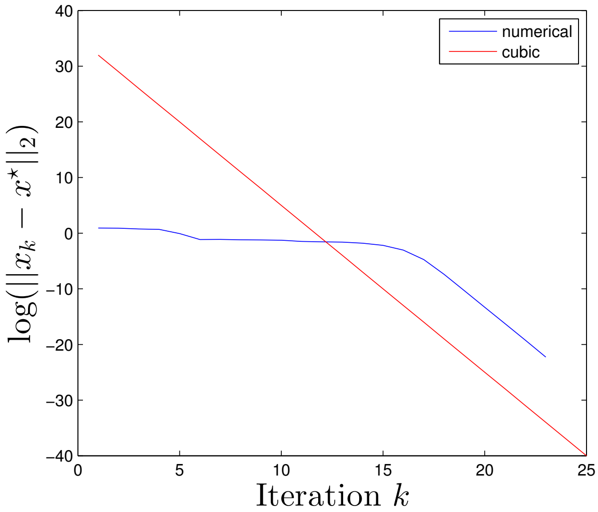


Figure 2: Order of convergence for Example 1.

Example 3: Bridge

In Figure 3 is seen the solution of a bridge construction. Even though the test case is somewhat bigger, the algorithm still runs fast, with error to $\sim 10^{-9}$ in 21 iterations.

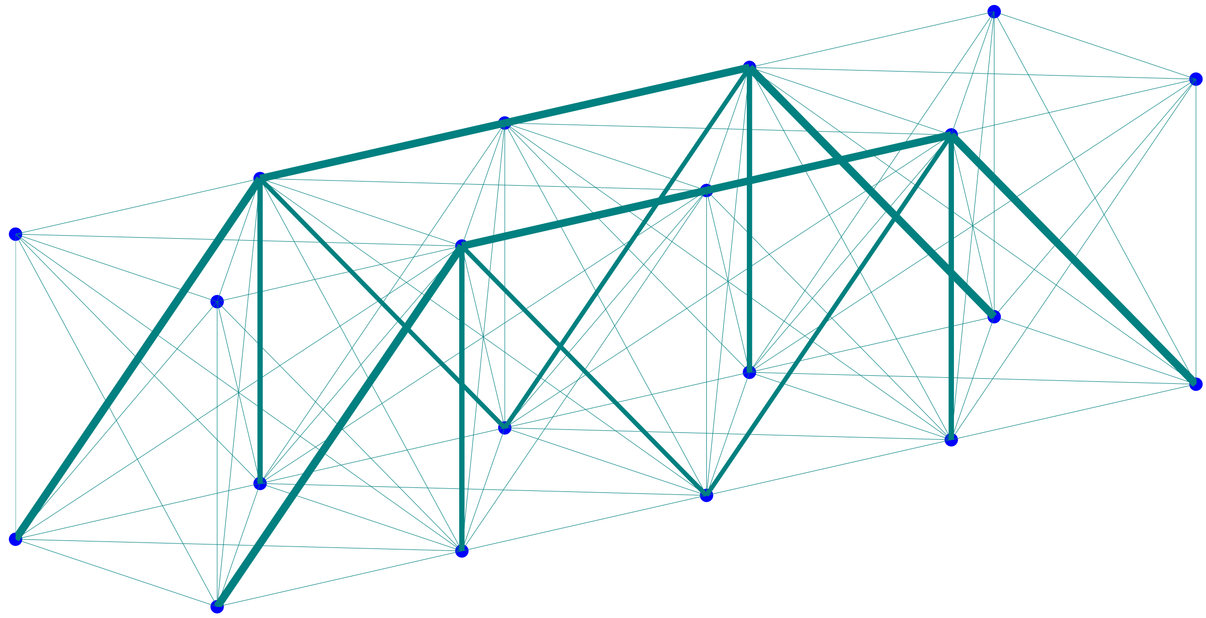


Figure 3: Truss example: Bridge. Numerical solution.

Example 4: Crane

At last we present an enormous test case of a crane. In Figure 4 can be seen both the construction and the numerical solution. As much as 97 nodes is used, and the external forces are split in two parts. First, at the hook in front, a vertical downforce is placed. To balance this, vertical downforce is placed at the other end of the horizontal beam. More specifically, the outer four nodes at the bottom of this part receives downforce. This tries to model the heavy load bricks which is usually seen on cranes.

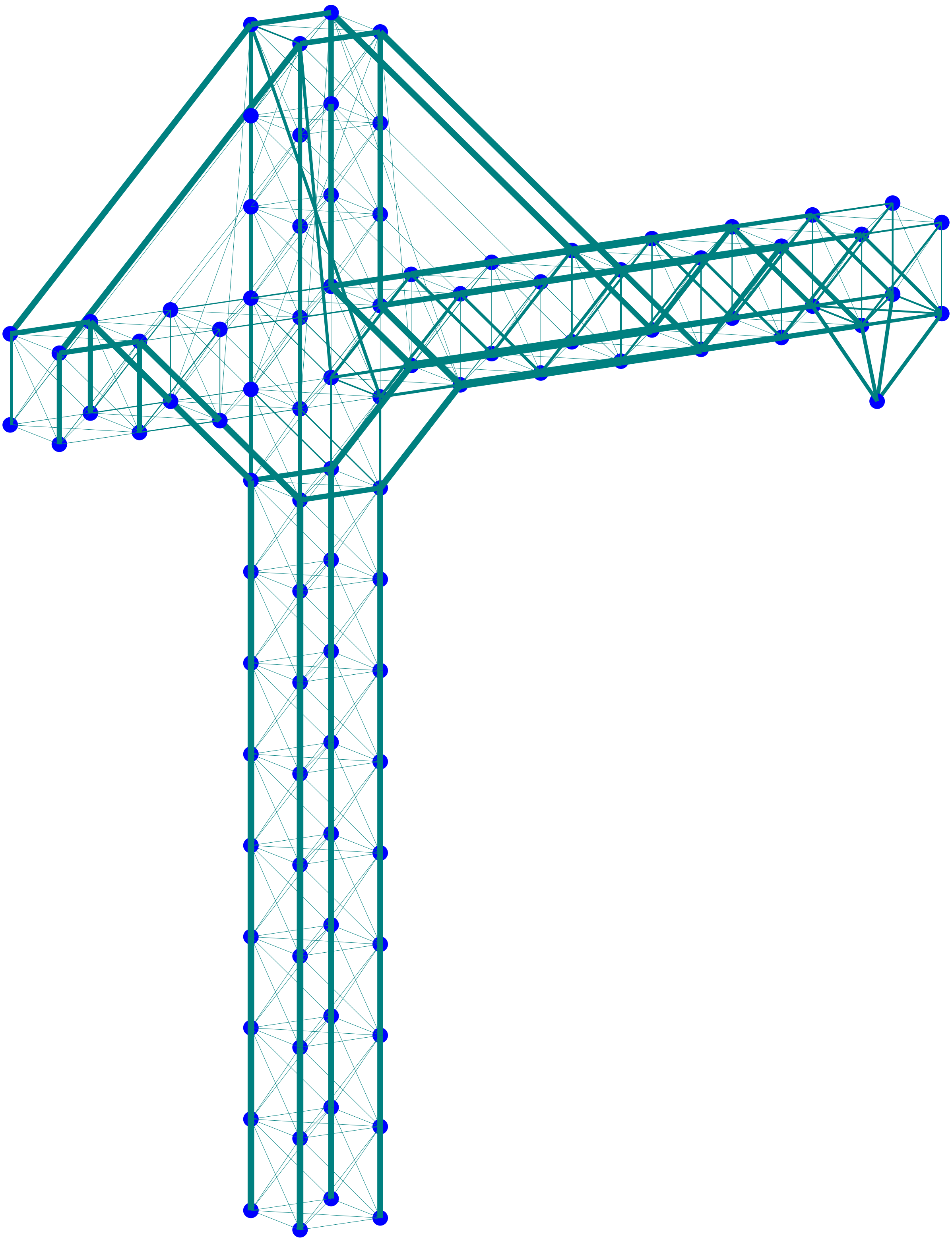


Figure 4: Truss example: Crane. Numerical solution.

References

[1] Francis Clarke, *Functional analysis, calculus of variations and optimal control*, Graduate Texts in Mathematics, vol. 264, Springer, London, 2013. MR 3026831

[2] Anton Evgrafov, *Lower semi-continuity, compactness and existence of solutions*, `http://www.math.ntnu.no/emner/TMA4180/2014v/Notes/note01.pdf`, 2014, [Online; accessed 7-March-2014].

[3] ———, *Optimality conditions for optimization over convex sets*, `http://www.math.ntnu.no/emner/TMA4180/2014v/Notes/note12.pdf`, 2014, [Online; accessed 17-March-2014].

[4] ———, *Optimization Theory Project/2014*, `http://www.math.ntnu.no/emner/TMA4180/2014v/Project/project.pdf`, 2014, [Online; accessed 7-March-2014].

[5] Harald E. Krogstad, *TMA 4180 Optimeringsteori KARUSH-KUHN-TUCKER THEOREM*, `http://www.math.ntnu.no/emner/TMA4180/2013v/HEKnotes/kkttheoremv2012.pdf`, 2013, [Online; accessed 24-March-2014].

[6] J. Nocedal and S. J. Wright, *Numerical optimization*, 2nd ed., Springer, New York, 2006.

[7] Wikipedia, *Conical combination*, `http://en.wikipedia.org/wiki/Conical_combination`, 2013, [Online; accessed 25-March-2014].

[8] ———, *Half-space (geometry)*, `http://en.wikipedia.org/wiki/Half-space_%28geometry%29`, 2013, [Online; accessed 10-March-2014].

[9] ———, *Matrix calculus*, `http://en.wikipedia.org/wiki/Matrix_calculus`, 2014, [Online; accessed 17-March-2014].

[10] ———, *Semi-continuity*, `http://en.wikipedia.org/wiki/Semi-continuity`, 2014, [Online; accessed 17-March-2014].