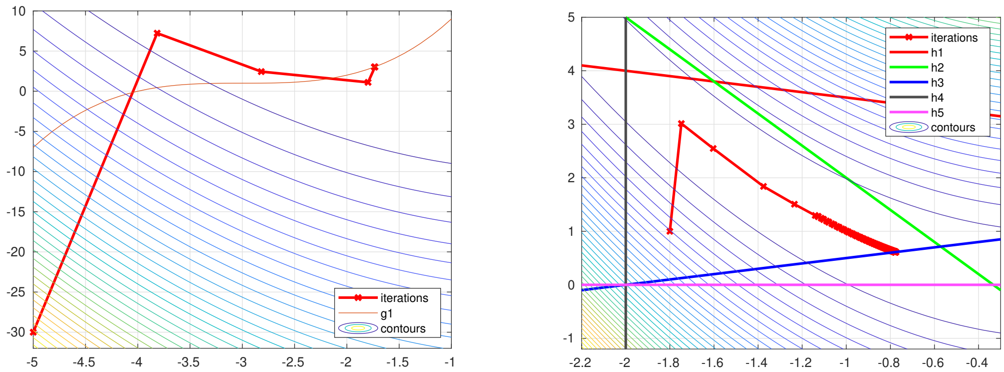
Prof. Dr.–Ing. habil. Thomas Meurer, Chair of Automatic Control



(a) Equality-constrained solution using (Newton-type) SQP.
(b) Inequality-constrained solution using unconstrained Newton line search (interior point).

Figure 2.1: Illustration of the solutions.

Exercise 1 (Local SQP method). Find the minimizer $\boldsymbol{x}^*$ of the equality-constrained problem given by

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2 \tag{2.1a}$$

subject to

$$g_1(\boldsymbol{x}) = (x_1 + 3)^3 - x_2 + 1 = 0. \tag{2.1b}$$

Before extending the template provided in [1] to solve (2.1), evaluate the following (preliminary) tasks by hand:

(a) Formulate the (general) KKT first order necessary optimality conditions $\nabla_{\boldsymbol{w}} l(\boldsymbol{w}) = \boldsymbol{0}$, where $\boldsymbol{w} = \left[\boldsymbol{x}^T \; \boldsymbol{\lambda}^T\right]^T$ using the Lagrangian $l(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x})$. What are the (matrix/vector) dimensions of the individual expressions?

(b) Note that the KKT conditions represent a system of $n + p$ (nonlinear) equations of the form

$$\boldsymbol{F}(\boldsymbol{w}^*) = \boldsymbol{0} \tag{2.2}$$

with $\boldsymbol{F}(\boldsymbol{w}) = \nabla_{\boldsymbol{w}} l(\boldsymbol{w})$. This nonlinear system of equations (2.2) can be solved numerically using the Newton (zero-finding) method, which generates a sequence of iterations $\boldsymbol{w}_k$ such that

$$\nabla_{\boldsymbol{w}} \boldsymbol{F}(\boldsymbol{w}_k) \left[\boldsymbol{w}_{k+1} - \boldsymbol{w}_k\right] = -\boldsymbol{F}(\boldsymbol{w}_k) \tag{2.3a}$$

$$\Leftrightarrow \boldsymbol{w}_{k+1} = \boldsymbol{w}_k - \left[\nabla_{\boldsymbol{w}} \boldsymbol{F}(\boldsymbol{w}_k)\right]^{-1} \boldsymbol{F}(\boldsymbol{w}_k) \tag{2.3b}$$

$$\Rightarrow \boldsymbol{w}_{k+1} = \boldsymbol{w}_k + \boldsymbol{r}_k, \tag{2.3c}$$

where $\boldsymbol{r}_k = -\left[\nabla_{\boldsymbol{w}} \boldsymbol{F}(\boldsymbol{w}_k)\right]^{-1} \boldsymbol{F}(\boldsymbol{w}_k)$ defines a step for the decision variables $\boldsymbol{x}$ and the Lagrange multipliers $\boldsymbol{\lambda}$. Calculate $\nabla_{\boldsymbol{w}} \boldsymbol{F}(\boldsymbol{w}_k)$ which is also known as the KKT matrix.
Hint: The KKT matrix can be calculated using $\nabla_{\boldsymbol{w}}^2 l(\boldsymbol{w}) = \frac{\partial}{\partial \boldsymbol{w}} \left(\nabla_{\boldsymbol{w}} l(\boldsymbol{w})\right) = \left[\frac{\partial(\nabla_{\boldsymbol{w}} l)}{\partial \boldsymbol{x}} \; \frac{\partial(\nabla_{\boldsymbol{w}} l)}{\partial \boldsymbol{\lambda}}\right]$,

see also [3]. Note that the KKT matrix contains the Hessian of the Lagrangian $L(\boldsymbol{w}) = \frac{\partial(\nabla_{\boldsymbol{x}} l)}{\partial \boldsymbol{x}}$

(c) Compare (2.3a) to the KKT conditions of an approximation of (2.1) using a (quadratic) Taylor series expansion of $f(\boldsymbol{x})$ and a (linear) Taylor series expansion of $\boldsymbol{g}(\boldsymbol{x})$ around some given iterate $\boldsymbol{x}_k$. Explain how/why the Newton method (2.3) is associated to the name *sequential quadratic programming*.

Implement the Newton method in `sqp` using the provided template [1] to solve (2.2) numerically. Hence,

(i) implement functions `rosen`, `gradRosen`, `hessRosen`, `equality`, `gradEquality`, and `hessEquality` in `oocLab2` which calculate the cost and constraint functions (2.1b) as well as their respective gradients and Hessians. Pass the function handles to the `handles` struct.

(ii) build the KKT matrix $\nabla_{\boldsymbol{w}} \boldsymbol{F}(\boldsymbol{w}_k)$ in `sqp`, calculate $\boldsymbol{r}_k$, and update the next iterate.

(i) graphically illustrate the successive iterations (after `sqp` finished) using `plot, ezcontour` and `fplot` so that you get something similar to fig. 2.1a

(iii) Analyze the quality of the solution using different starting values $\boldsymbol{x}_{\text{start}}, \boldsymbol{\lambda}_{\text{start}}$.

(iv) Extend your code to implement the damped BFGS update (Remark 3.4 in the lecture notes) to approximate the Hessian of the Lagrangian in the KKT matrix.

(iv) Does the damped BFGS update improve the convergence behavior of arbitrary starting values $\boldsymbol{w}_{\text{start}}$ compared to using the analytical Hessian? Also compare the total number of iterations.

Exercise 2 (Interior point method). Find the minimizer of the inequality-constrained problem given by

$$\min_{\boldsymbol{x} \in \mathbb{R}^2} f(\boldsymbol{x}) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2 \tag{2.4a}$$

subject to

$$h_1(\boldsymbol{x}) = x_2 + \frac{1}{2}x_1 - 3 \leq 0, \tag{2.4b}$$
$$h_2(\boldsymbol{x}) = x_2 + 3x_1 + 1 \leq 0, \tag{2.4c}$$
$$h_3(\boldsymbol{x}) = -x_2 + \frac{1}{2}x_1 + 1 \leq 0, \tag{2.4d}$$
$$h_4(\boldsymbol{x}) = -x_1 - 2 \leq 0, \tag{2.4e}$$
$$h_5(\boldsymbol{x}) = -x_2 \leq 0. \tag{2.4f}$$

Subsequently, extend your code from Computer Lab 1 using the Newton line search method by implementing Algorithm 7 from the lecture notes [2] using the barrier function

$$B(\boldsymbol{x}) = -\sum_{l=1}^{q} \frac{1}{h_l(\boldsymbol{x})}. \tag{2.5}$$

In the following,

(i) implement the functions `barrier,gradBarrier,hesseBarrier`. Note that the Barrier function can be viewed as $B(\boldsymbol{x}) = \Phi(\boldsymbol{h}(\boldsymbol{x}))$ so that you can apply the chain rule to get $\nabla_{\boldsymbol{x}} B(\cdot), \nabla^2_{\boldsymbol{x}} B(\cdot)$.

(ii) extend your code so that cost function, gradient and Hessian values from your previous implementation are altered only if inequalities are present.

(iii) adapt the stopping criterion so that a sequence of unconstrained problems is solved with increasing c_k.

(iv) since interior point methods require the starting point to be inside the feasible region determine a feasible starting point such that $h_i(\boldsymbol{x}_{\text{start}}) \leq 0, \quad \forall i = 1, \dots, q$.

(v) graphically illustrate the successive iterations (after the interior point `lineSearch` finished) using `plot, ezcontour` and `fplot` so that you get something similar to fig. 2.1b

(vi) which inequalities are active at the solution? What are the corresponding Lagrange multipliers $\boldsymbol{\mu}$?
Hint: To get an estimate for the Lagrange multipliers, compare the necessary optimality conditions for the auxiliary unconstrained problem $\min_{\boldsymbol{x} \in \mathbb{R}^n} \tilde{f}(\boldsymbol{x}) = f(\boldsymbol{x}) + 1/cB(\boldsymbol{x})$ with the general KKT first order necessary optimality conditions for inequality-constrained problems.

References

[1] J. Andrej, D. Siebelts, and S. Helling. *oocLab2*. `https://cau-git.rz.uni-kiel.de/ACON/opt/optimization-and-optimal-control`. 2020 (cit. on pp. 1, 2).

[2] T. Meurer. *Optimization and Optimal Control WS 20/21*. `https://www.control.tf.uni-kiel.de/en/teaching/winter-term/optimization-and-optimal-control/fileadmin/opt_ws2021_full`. 2020 (cit. on p. 2).

[3] K. B. Petersen and M. Pedersen. *The Matrix Cookbook*. `https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf`. 2012 (cit. on p. 2).