Worked solutions for Sample Exam 2
“Optimization in Systems and Control” (SC4091)

QUESTION 1: Optimization methods I

Please note that for some questions more than one answer might be correct. However, below only one answer is listed. Furthermore, the footnotes are for further clarification only and are not considered to be a required part of the answers.
The various εs appearing in the stopping criteria below are all assumed to be small positive numbers.

Answers for Question 1

P1. (a) The objective function of this optimization problem is convex in its argument x since it is the sum of three convex functions:

* $|2x_1 - x_2 + x_3 - 3x_4|$ is convex since the absolute value function is convex and since a convex function of an affine function is also convex,
* $-6\log(5+x_1+x_2+x_3)$ is convex since $-\log(x)$ (and thus also $-6\log(x)$) is a convex function, and since a convex function of an affine function is also convex,
* x_4^2 is convex.

The constraint is also convex since $g(x) = x_1^2 + x_2^4 + x_3^2 + x_4^4 - 1$ is the sum of several convex functions. Hence, the optimization problem is convex. This implies that a suited optimization algorithm is the ellipsoid method (M12).

(b) For the ellipsoid method a suited stopping criterion is

$$|f(x_k) - f(x^*)| \leq \varepsilon_f, \quad \|x_k - x^*\|_2 \leq \varepsilon_x \quad \text{and} \quad g(x_k) \leq \varepsilon_g \ ,$$

where x_k is the current iteration point, x^* is the (yet unknown) optimum of this optimization problem, f denotes the objective function, and the constraint function g is as defined above.

P2. (a) This maximization problem can be transformed into a minimization problem by considering the opposite of the objective function. This results in

$$\min_{x\in\mathbb{R}^4} -\frac{x_1 + x_2^2 - x_1x_2 - x_3^3 + 2x_4}{1 + 5\sqrt{x_1^4 + x_2^6 + x_3^8 + x_4^4}} \ . \tag{1}$$

This problem is an unconstrained optimization problem with a non-convex objective function (e.g., due to the presence of $-x_3^3$). Hence, a multi-start method is required. The gradient and the Hessian of the objective function can be computed analytically. This implies that a suited optimization algorithm is *multi-start* Levenberg-Marquardt[1] (M6).

(b) The most appropriate stopping criterion is

$$\|\nabla f(x_k)\|_2 \leq \varepsilon \quad ,$$

where x_k is the current iteration point and f is the objective function of the optimization problem (1).

P3. (a) This maximization problem can be transformed into a minimization problem by considering the opposite of the objective function. The objective function of the resulting minimization problem is non-convex (e.g., due to the cos term). Hence, we have a nonlinear, non-convex optimization problem and therefore a multi-start method is required.
The constraints of the problem can also be simplified somewhat: Since log is a non-decreasing function the first constraint can be recast as follows:

$$\exp(-2) \leq 1 + 3x_1^2 + 2x_2^2 + 8x_3^2 \leq \exp(3)$$

The second constraint can be rewritten as $-4 \leq x_i \leq 4$ for $i = 1, 2, \ldots, 5$.
The gradient and Hessian of the objective function and the Jacobian of the constraints can be computed analytically. This implies that a suited optimization algorithm is *multi-start* sequential quadratic programming (M3).

(b) The most appropriate stopping criterion is[2]: there exists a μ_k such that

$$\begin{aligned}
&\| \nabla f(x_k) + \nabla g(x_k)\, \mu_k \|_2 \leq \varepsilon_1 \\
&| \mu_k^{\mathrm{T}} g(x_k) | \leq \varepsilon_2 \\
&\mu_k \geq -\varepsilon_3 \\
&g(x_k) \leq \varepsilon_4
\end{aligned}$$

where x_k is the current iteration point, f is the objective of the minimization problem, and g is the inequality constraint function (written in the form $g(x) \leq 0$) of the simplified optimization problem.

P4. (a) By introducing a dummy variable t and adding extra constraints the problem

[1]The Levenberg-Marquardt algorithm is to be preferred above the BFGS quasi-Newton method since the Levenberg-Marquardt algorithm uses the exact Hessian (i.e., 2nd-order information) whenever possible, whereas the BFGS method uses an approximation of the Hessian based on function values and gradients (i.e., 1st-order information).

[2]There are no equality constraints; so the equality constraint function h does not appear here.

$\min_{x\in\mathbb{R}^4} \max(|x_1|, 2|x_2|, 3|x_3|, 4|x_4|)$ can be recast as

$$\begin{aligned} &\min_{t\in\mathbb{R}, x\in\mathbb{R}^4} \quad t \\ &\text{subject to} \quad -t \leq x_1 \leq t \\ &\qquad\qquad\quad -t \leq 2x_2 \leq t \\ &\qquad\qquad\quad -t \leq 3x_3 \leq t \\ &\qquad\qquad\quad -t \leq 4x_4 \leq t \quad . \end{aligned} \tag{2}$$

This is a linear programming problem. The constraints of the original problem can be recast as

$$x_1 + x_2 + x_3 + x_4 \leq 9$$

and

$$-6 \leq x_1 + x_2 - x_3 - 6x_4 \leq 6 \quad .$$

Adding these constraints to the problem (2) still results in a linear programming problem. So a suited optimization algorithm is the simplex algorithm (M1).

(b) Since the simplex algorithm finds the optimal solution in a finite number of steps, no stopping criterion is required[3].

P5. (a) This problem is an unconstrained nonlinear least-squares problem. Since the problem is non-convex, a multi-start method is required. The gradients and Hessians of the components of the error function can be computed analytically. Hence, an appropriate method is the *multi-start* Gauss-Newton least-squares algorithm (M5).

(b) An appropriate stopping criterion is

$$\|\nabla e(\theta_k)\|_2 \leq \varepsilon \quad ,$$

where e denotes the error function (i.e., the optimization problem is of the form $\min_{\theta\in\mathbb{R}^7} \|e(\theta)\|_2$).

P6. (a) The objective function is a quadratic function. It is convex since it can be written as $\frac{1}{2}x^{\mathrm{T}}Hx + cx + d$ with

$$H = \begin{bmatrix} 8 & 4 & 6 \\ 4 & 10 & 2 \\ 6 & 2 & 6 \end{bmatrix}$$

and with H positive definite (since its leading principal minors all have positive determinants).

The first constraint of the given optimization problem is convex[4] since $g(x) = x_1^2 + 2x_2^2 +$

[3] However, in practice a maximum number of iterations is usually specified.

[4] This constraint cannot be further simplified into a linear constraint. Hence, the problem is not a quadratic programming problem.

$6x_3^2 - 4$ is a convex function. Indeed, the scalar functions $x_i \mapsto x_i^2$ and $x \mapsto -4$ are convex, and a weighted sum of convex function with positive weights is also convex.
The second constraint of the given optimization problem can be recast as two linear — and thus convex — inequality constraints.
So the problem is a convex optimization problem. This implies that a suited optimization algorithm is the ellipsoid method (M12).

(b) For the ellipsoid method a suited stopping criterion is

$$|f(x_k) - f(x^*)| \leq \varepsilon_f, \quad \|x_k - x^*\|_2 \leq \varepsilon_x \quad \text{and} \quad g(x_k) \leq \varepsilon_g \ ,$$

where x_k is the current iteration point, x^* is the (yet unknown) optimum of this optimization problem, f is the objective function, and g contains the inequality constraints (in the form $g(x) \leq 0$).

P7. (a) The objective function is nonlinear and non-convex. Hence, we have a nonlinear, non-convex optimization problem and therefore a multi-start optimization algorithm is required. The constraints are linear. Computing an analytic expression for the gradient of the objective function is not easy. Furthermore, evaluating the function requires numerical integration, which is computationally expensive. This implies that it is not recommended to compute the gradient of the objective function analytically or numerically. As a consequence, a gradient-free optimization method is recommended. By using a penalty or barrier function we can transform the problem into an unconstrained optimization problem. The most suited algorithm for this constrained optimization problem is a *multi-start* penalty function method + line search method with Powell directions and line minimization using parabolic interpolation (M11).

(b) A suited stopping criterion is[5]

$$|f(x_{k+1}) - f(x_k)| \leq \varepsilon_f \quad \text{and} \quad \|x_{k+1} - x_k\|_2 \leq \varepsilon_x \ ,$$

where f is the objective function of the unconstrained problem (so it also includes the penalty term).

P8. (a) Since the objective function is nonlinear and non-convex (e.g., due to the x_1^3 term), we have a nonlinear, non-convex optimization problem. This implies that a multi-start method is required. It is not possible to eliminate one of the variables from the equality constraint. Therefore, the Lagrange method can be used to transform the problem into an unconstrained problem. The gradient and Hessian of the objective function of the resulting unconstrained optimization problem can be computed analytically. This implies that a suited optimization algorithm is a *multi-start* Lagrange method + Levenberg-Marquardt algorithm (M9).

[5]We cannot use the gradient here as we have argued in part (a) that the gradient should not be used.

(b) The most appropriate stopping criterion is

$$\begin{aligned}&\|\nabla f(x_k) + \nabla h(x_k)\lambda_k\|_2 \leq \varepsilon_1\\&\|h(x_k)\|_2 \leq \varepsilon_2 \ ,\end{aligned}$$

where f and h are respectively the objective function and the equality constraint function of the original constrained problem (with the equality constraint written in the form $h(x) = 0$).

P9. (a) First denote the quadratic argument inside the brackets by g. So $g(x) = 7x_1^2 + 2x_2^2 + x_3^2 - 4x_1x_2 + 2x_2x_3 + 1$. Note that g can also be rewritten as a sum of squares: $g(x) = 3x_1^2 + (2x_1 - x_2)^2 + (x_2 + x_3)^2 + 1$. So g is a nonnegative function. Now, since the function h defined by $h(y) = y^{\frac{1}{3}}$ is a non-decreasing function for $y \geq 0$, we can simplify the problem to the minimization of g subject to the given constraints. The (convex) constraint $\|x\|_1 \leq 1$ or equivalently $|x_1| + |x_2| + |x_3| \leq 1$ can be recast as a collection of $2^3 = 8$ linear constraints:

$$\begin{aligned}x_1 + x_2 + x_3 &\leq 1\\x_1 + x_2 - x_3 &\leq 1\\x_1 - x_2 + x_3 &\leq 1\\x_1 - x_2 - x_3 &\leq 1\\-x_1 + x_2 + x_3 &\leq 1\\-x_1 - x_2 + x_3 &\leq 1\\-x_1 + x_2 - x_3 &\leq 1\\-x_1 - x_2 - x_3 &\leq 1 \ .\end{aligned}$$

This implies that we have a (convex) quadratic optimization problem. Therefore, a suited optimization algorithm is the modified simplex algorithm for quadratic programming (M2).

(b) Since the modified simplex algorithm finds the optimal solution in a finite number of steps, no stopping criterion is required.[6]

P10. (a) Since the function g defined by $g(y) = 2^y$ is a non-decreasing function, we can simplify the problem to the minimization of $|x_1| + |x_2| + |x_3|$ subject to the given constraints. It is moreover easy to verify that the constraints are linear: the first constraint can be recast as $x_1 - x_2 + 8 \geq 5$, $9 - 2x_1 + 3x_2 \geq 5$, and the second constraint can be recast as $(x_1 + x_2 - x_3 - 6)^2 \leq \operatorname{acosh}(100)$ or as $-\sqrt{\operatorname{acosh}(100)} \leq x_1 + x_2 - x_3 - 6 \leq \sqrt{\operatorname{acosh}(100)}$, where acosh denotes the inverse of cosh. By introducing 3 dummy variables α_1, α_2, and α_3, considering the objective function $\alpha_1 + \alpha_2 + \alpha_3$ and adding the constraints $\alpha_i \geq |x_i|$ for $i = 1,2,3$ or equivalently, $\alpha_i \geq x_i$ and $\alpha_i \geq -x_i$ for $i = 1,2,3$ we end up with a linear optimization problem. Therefore, a suited optimization algorithm is the simplex algorithm for linear programming (M1).

[6]However, in practice a maximum number of iterations is usually specified.

(b) Since the simplex algorithm finds the optimal solution in a finite number of steps, no stopping criterion is required.

Answers for Question 2

The gradient of the objective function f is given by

$$\nabla f(x) = \begin{bmatrix} 8x_1 + 4x_2 + 6x_3 - 1 \\ 4x_1 + 10x_2 + 2x_3 + 1 \\ 6x_1 + 2x_2 + 6x_3 \end{bmatrix} .$$

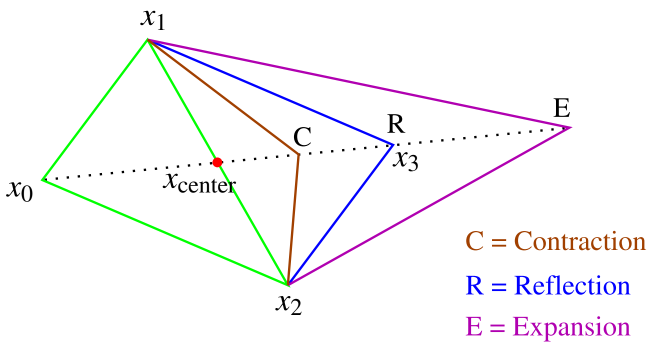
The Hessian of f is given by

$$H(x) = \begin{bmatrix} 8 & 4 & 6 \\ 4 & 10 & 2 \\ 6 & 2 & 6 \end{bmatrix} .$$

QUESTION 2: Optimization methods II

Answer for Question a)

[Here we essentially expect a summary of Section 4.3 of the lecture notes including figures like

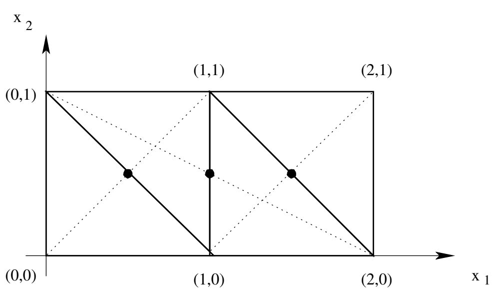


and like Figure 4.5, and an explanation of these figures. Be sure to mention and/or define the following elements: simplex, $n+1$ points in $\mathbb{R}^n$, *point* reflection of the point with largest function value around center of the other points, $x_{\text{new}} = x_{\text{center}} + d$ with $d = x_{\text{center}} - x_{\text{largest function value}}$, iteration results in non-increasing sequence of function values, stop if reflection does not yield a smaller function value any longer, extensions: contraction and expansion, formula: $x_{\text{new}} = x_{\text{center}} + \alpha d$ with $\alpha < 1$ for contraction and $\alpha > 1$ for expansion.]

Answer[7] for Question b)

The initial simplex is defined by the points $(0,0)$, $(0,1)$, and $(1,0)$. We have $f(0,0) = 0$, $f(0,1) = -1$, $f(1,0) = -3$. So $(0,0)$ is the point with the largest function value. Hence, we reflect $(0,0)$ around the center of $(0,1)$, and $(1,0)$, i.e., around $(\frac{1}{2}, \frac{1}{2})$. This yields the point $(1,1)$ and a new simplex defined by the points $(0,1)$, $(1,0)$, $(1,1)$.
We have $f(1,1) = -5$. So $(0,1)$ is now the point with the largest function value. Hence, we reflect $(0,1)$ around the center of $(1,0)$, and $(1,1)$, i.e., around $(1, \frac{1}{2})$. This yields the point $(2,0)$ and a new simplex defined by the points $(1,0)$, $(1,1)$ and $(2,0)$.
We have $f(2,0) = -4$. Now $(1,0)$ is the point with the largest function value. Hence, we reflect $(1,0)$ around the center of $(1,1)$, and $(2,0)$, i.e., around $(\frac{3}{2}, \frac{1}{2})$. This yields the point $(2,1)$ and a new simplex defined by the points $(1,1)$, $(2,0)$ and $(2,1)$. We have $f(2,1) = -7$.

[7] For completeness we also provide here a figure illustrating the various steps:



QUESTION 3: Controller design

Answer for Question a)

We have

$$\begin{aligned} y(k) &= d(k) + P(q)\big(d_{\mathrm{p}}(k) - K(q)\big(d_{\mathrm{s}}(k) + y(k)\big)\big) \\ &= d(k) + P(q)d_{\mathrm{p}}(k) - P(q)K(q)d_{\mathrm{s}}(k) - P(q)K(q)y(k) \ . \end{aligned}$$

Hence,

$$y(k) = \big(1 + P(q)K(q)\big)^{-1}\big(d(k) + P(q)d_{\mathrm{p}}(k) - P(q)K(q)d_{\mathrm{s}}(k)\big) \ .$$

We have

$$\begin{aligned} v(k) &= K(q)\big(d_{\mathrm{s}}(k) + d(k) + P(q)\big(d_{\mathrm{p}}(k) - v(k)\big)\big) \\ &= K(q)d_{\mathrm{s}}(k) + K(q)d(k) + K(q)P(q)d_{\mathrm{p}}(k) - K(q)P(q)v(k) \ . \end{aligned}$$

Hence,

$$v(k) = \big(1 + P(q)K(q)\big)^{-1}\big(K(q)d(k) + K(q)P(q)d_{\mathrm{p}}(k) + K(q)d_{\mathrm{s}}(k)\big) \ .$$

So

$$M(q) = \big(1 + P(q)K(q)\big)^{-1} \begin{bmatrix} 1 & P(q) & -P(q)K(q) \\ K(q) & K(q)P(q) & K(q) \end{bmatrix} \ .$$

Answer for Question b)

First we consider the expression $(1+PK)^{-1}$ (where for the sake of simplicity of notation we have dropped the q operator as argument). Substituting $K = Q(1-PQ)^{-1}$ results in $(1+PK)^{-1} = \big(1 + \frac{PQ}{1-PQ}\big)^{-1} = \big((1-PQ)^{-1}\big)^{-1} = 1 - PQ$. So

$$\begin{aligned} M &= (1-PQ)\begin{bmatrix} 1 & P & -PQ(1-PQ)^{-1} \\ Q(1-PQ)^{-1} & Q(1-PQ)^{-1}P & Q(1-PQ)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} 1-PQ & (1-PQ)P & -PQ \\ Q & QP & Q \end{bmatrix} \ . \end{aligned}$$

So the entries of M consist of sums and products of stable transfer functions (i.e., P and Q) and since the set of stable transfer functions is closed under addition and multiplication, this implies that all entries of M are stable.

Answer for Question c)

We have $M_{11} = 1 - PQ$. So in order to prove that the constraint $\|M_{11}\|_\infty \leq 1$ is closed-loop convex, we have to prove that for all rational stable transfer functions Q_1 and Q_2 such that $\|1 - PQ_1\|_\infty \leq 1$ and $\|1 - PQ_2\|_\infty \leq 1$ and for all $\lambda \in [0,1]$ we have $\|1 - P\big((1-\lambda)Q_1 + \lambda Q_2\big)\|_\infty \leq 1$. Since

$$
\begin{aligned}
\|1 - P((1-\lambda)Q_1 + \lambda Q_2)\|_\infty &= \|1 - \lambda + \lambda - P\big((1-\lambda)Q_1 + \lambda Q_2\big)\|_\infty \\
&= \|(1-\lambda) + \lambda - (1-\lambda)PQ_1 - \lambda PQ_2)\|_\infty \\
&= \|(1-\lambda)(1 - PQ_1) + \lambda(1 - PQ_2)\|_\infty \\
&\leq \|(1-\lambda)(1 - PQ_1)\|_\infty + \|\lambda(1 - PQ_2)\|_\infty \quad \text{(by the triangle inequality)} \\
&\leq |1-\lambda| \underbrace{\|1 - PQ_1\|_\infty}_{\leq 1} + |\lambda| \underbrace{\|1 - PQ_2\|_\infty}_{\leq 1} \\
&\leq |1-\lambda| + |\lambda| \\
&\leq 1 - \lambda + \lambda \qquad \text{(since } 0 \leq \lambda \leq 1\text{)} \\
&\leq 1 \ ,
\end{aligned}
$$

the constraint $\|M_{11}\|_\infty \leq 1$ is indeed closed-loop convex.