Worked solutions for Exam of October 2013
“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.

Answers

P1. (a) We first write the maximization problem as a minimization problem, also taking into account that $\exp(\cdot)$ is a non-decreasing function, which means that we can also minimize its argument instead. This yields

$$\min_{x \in \mathbb{R}^3} x_1^2 + 2x_1x_2 + x_2^2 + x_3^2 - x_1 - 3x_2 + 4x_3 \quad . \tag{1}$$

The first constraint of the given problem contains a 1-norm and it can be expanded as follows:

$$|x_1| + |x_2| + |x_3| \leqslant 1 \quad ,$$

or equivalently

$$\begin{aligned} x_1 + x_2 + x_3 &\leqslant 1 \\ x_1 + x_2 - x_3 &\leqslant 1 \\ x_1 - x_2 + x_3 &\leqslant 1 \\ x_1 - x_2 - x_3 &\leqslant 1 \\ &\vdots \\ -x_1 - x_2 - x_3 &\leqslant 1 \end{aligned}$$

(8 constraints in total). Note that these are linear[1] constraints.
Since the function $\log(\cdot)$ in the second constraint is non-decreasing and since the first constraint ensures that the argument of the $\log(\cdot)$ function is always strictly positive, we can recast the constraint as

$$4 + x_1 + x_2 - x_3 \leqslant e \quad ,$$

which is a linear constraint.
It is easy to verify that the (simplified) objective function as given in (1) is convex (as it can be written as $(x_1 + x_2)^2 + x_3^2 - x_1 - 3x_2 + 4x_3$, which is a sum of convex functions[2]).

[1] Or better: affine.
[2] Alternatively, it can be verified that the Hessian of this objective function is positive semi-definite.

Hence, we have a convex quadratic programming (QP) problem. Since the modified simplex algorithm is not listed, we select the next best option in the list of algorithms, i.e., the interior point algorithm (M11).

(b) The corresponding stopping criterion is[3]

$$|f_{\mathrm{QP}}(x_k) - f_{\mathrm{QP}}(x^*)| \leqslant \varepsilon \ ,$$

where x_k is the current iteration point, f_{QP} is the objective function of the QP problem (cf. (1)), and ε is a small positive number.

P2. (a) Although the objective function is convex, the constraint is not: in standard form the constraint is $4 - \|x\|_2 \leqslant 0$, and $-\|x\|_2$ is a concave function, not a convex one. Hence, we have a nonlinear, non-convex optimization problem with inequality constraints. The gradient and Hessian can be computed easily. Therefore, the most appropriate method from the given list of algorithms is a *multi-start* penalty function method with steepest descent line minimization (M9).

(b) The corresponding stopping criterion is: there exists a (scalar) μ_k such that

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

where x_k is the current iteration point, f is the objective of the minimization problem, g is the inequality constraint function (when the constraint is written in the form $g(x) \leqslant 0$), and ε_1, ε_2, ε_3, ε_4 are small positive numbers.

P3. (a) Since $(\cdot)^3$ is non-decreasing, we can also minimize its argument. By introducing dummy variables $\alpha_1, \ldots, \alpha_5$ with $\alpha_i \geqslant |x_i|$ and subsequently minimizing each α_i, we can write the given optimization problem as

$$\begin{aligned}
\min_{\alpha, x \in \mathbb{R}^5} \ & 1 + \alpha_1 + 3\alpha_2 + \alpha_3 + 2\alpha_4 + 3\alpha_5 \\
\text{s.t.} \ & -\alpha_i \leqslant x_i \leqslant \alpha_i \qquad \text{for } i = 1, \ldots, 5 \\
& -3 \leqslant 5 + 3x_1 + x_2 - x_3 + x_4 - x_5 \leqslant 3 \\
& 3 + 2x_1 + 3x_2 \geqslant 2 \\
& 4x_3 + 3x_4 \geqslant 2 \\
& 2x_5 + 8 \geqslant 2 \ .
\end{aligned}$$

[3]Note that the points returned by the interior point algorithm are always feasible by construction; so we do *not* have to check explicitly where the constraint violation is small enough.

So we have a linear programming problem, and the most suited method is the simplex algorithm for linear programming (M1).

(b) Since the simplex algorithm finds the optimal solution in a finite number of steps, no stopping criterion is required[4].

P4. (a) The first term of the objective function is of the form $1/y^6$ with $y \geqslant 1$ (due to the first constraint). Since the domain of this function is convex, and since its second derivative (i.e., $(-6)(-7)1/y^8$) is always positive for $y \geqslant 1$, the function $f_1 : [1,\infty) \to \mathbb{R} : y \mapsto 1/y^6$ is convex. Since replacing the argument of a convex function by an affine expression preserves convexity, the first term of the objective function is convex.
As for the second term of the objective function, first note that the function $f_2 : y \mapsto \cosh|y|$ is convex. Indeed, $|y|$ is convex (and nonnegative), and $\cosh(\cdot)$ is also convex, and non-decreasing for nonnegative arguments. If we then replace y by the affine expression $x_1 + 4x_2 + 5x_3 - 100$, convexity is preserved.
Since the overall objective function is a linear combination (with positive weights) of two convex functions, it is also convex.
The first constraint is linear and thus convex.
Since $\log_{10}(\cdot)$ is a non-decreasing function, the second constraint can be rewritten as

$$(x_1-2)^2 + (x_2-3)^2 + (x_3-4)^2 \leqslant 100 \ .$$

Since $f_3 : y \mapsto y^2$ is convex and since convexity is preserved by considering an affine argument and by addition, this constraint is also convex.
So we have a convex optimization problem, and therefore we select the interior point algorithm (M11).

(b) The corresponding stopping criterion is[5]

$$|f(x_k) - f(x^*)| \leqslant \varepsilon \ ,$$

where x_k is the current iteration point, f is the objective function, and ε is a small positive number.

P5. (a) As we have a nonlinear equality constraint, the given problem is a nonlinear non-convex optimization problem with equality constraints. It is not possible to use the constraint to eliminate one of the variables. The gradient and Hessian of the objective function and of the constraint function can be computed easily. Hence, the most appropriate method is a *multi-start* Lagrange method with the Levenberg-Marquardt direction for the line minimization (M6).

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

[5] The points returned by the interior point algorithm are always feasible by construction; so we do *not* have to check explicitly where the constraint violation is small enough.

(b) The corresponding stopping criterion is: there exists a λ_k such that

$$\| \nabla f(x_k) + \nabla h(x_k)\, \lambda_k \|_2 \leqslant \varepsilon_1$$
$$\|h(x_k)\|_2 \leqslant \varepsilon_2$$

where x_k is the current iteration point, f is the objective of the minimization problem, h is the equality constraint function (when the constraint is written in the form $h(x) = 0$), and ε_1, ε_2 are small positive numbers.

P6. (a) By introducing a scalar dummy variable t such that $t \leqslant \min(|4x_1 - 3x_2 + 8x_3 - 5|, |-2x_1 + 7x_2 - x_3 + 1|)$ and subsequently maximizing t, and by taking into account that a constraint of the form $t \leqslant |\alpha|$ with $\alpha \in \mathbb{R}$ is equivalent to $t \leqslant \alpha$ **or** $t \leqslant -\alpha$, the maximization of the objective function can be rewritten as

$$\begin{aligned}
\max_{x\in\mathbb{R}^3, t\in\mathbb{R}} \quad & t \\
\text{s.t.} \quad & \big(t \leqslant 4x_1 - 3x_2 + 8x_3 - 5 \\
& \quad \textbf{or } t \leqslant -(4x_1 - 3x_2 + 8x_3 - 5) \big) \ \textbf{ and} \\
& \big(t \leqslant -2x_1 + 7x_2 - x_3 + 1 \\
& \quad \textbf{or } t \leqslant -(-2x_1 + 7x_2 - x_3 + 1) \big) \ .
\end{aligned}$$

So we get a total of 4 problems with a linear objective function and linear constraints. Taking into account that $f_1 : y \mapsto 2^y$ is a non-decreasing function and that $\|x\|_\infty = \max_i |x_i|$, the constraints of the original optimization problem can be rewritten as

$$\begin{aligned}
& 7x_1 - 2x_2 + 5x_3 \leqslant 6 \\
& -3 \leqslant x_i \leqslant 3 \qquad \text{for } i = 1, 2, 3\,,
\end{aligned}$$

which are also linear.
Hence, we have 4 times a linear programming problem and we have to take the overall maximum of the obtained solutions. The most suited method is the simplex algorithm for linear programming (M1).

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

P7. (a) By taking into account that $f_1 : y \mapsto \arctan(y)$ is a non-decreasing function, and by simplifying the constraints, we obtain the following quadratic programming (QP) problem:

$$\begin{aligned}
\min_{x\in\mathbb{R}^3} \ & 4x_1^2 + 2x_1x_2 + 2x_1x_3 + 4x_2^2 + 2x_2x_3 + 4x_3^2 - x_1 + x_2 - x_3 \\
\text{s.t.} \ & -5 \leqslant 3x_1 - x_2 + 6x_3 \leqslant 5 \\
& -9 \leqslant x_1 + x_2 + x_3 \leqslant 9 \ .
\end{aligned}$$

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

By investigating the Hessian of the objective function of this QP problem, which is given by

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

and which is positive definite (as all leading principal minors are positive), or by noticing that the objective function of the QP problem can also be rewritten as $(x_1 + x_2 + x_3)^2 + 3x_1^2 + 3x_2^2 + 3x_3^2 + (-x_1 + x_2 - x_3)$, which is a sum of convex functions, it can be asserted that this objective function is convex (and quadratic). Since the modified simplex algorithm is not listed, we select the next best option in the list of algorithms, i.e., the interior point algorithm (M11).

(b) The corresponding stopping criterion is[7]

$$|f_{\mathrm{QP}}(x_k) - f_{\mathrm{QP}}(x^*)| \leqslant \varepsilon$$

where x_k is the current iteration point, f_{QP} is the objective function of the QP problem, and ε is a small positive number.

P8. (a) The maximization problem can be recast as a minimization problem: $\min_{v \in \mathbb{R}^{10}} (-\|G\|_\infty)$. Clearly, the objective function is not convex. Since the ∞-norm of a transfer function has to be computed numerically, and since the number of variables is large, it will be very time-consuming to compute the gradient and the Hessian of the objective function numerically. Therefore, we should select a gradient-free unconstrained[8] optimization method from the list of algorithms. Hence, we could select *multi-start* simulated annealing (M10) or a *multi-run* genetic algorithm (M12).

(b) Corresponding suitable stopping criteria are that the temperature in the simulated annealing algorithm is less than some threshold ($T \leqslant T_{\mathrm{final}}$), or — for the genetic algorithm — that the maximum number of generations has been reached.

P9. (a) The objective function is non-convex due to the terms $3x_1x_3$ and $2x_4x_5$ (which cannot be absorbed into a square of an affine function as could be done in e.g. Problem P7 above). Since $\|x\|_\infty = \max_i |x_i|$ the constraints can be rewritten as linear constraints:

$$\begin{aligned} &-7 \leqslant x_1 + 2x_2 + 8x_3 - 9x_4 + 8x_5 \leqslant 7 \\ &-x_1 + 3x_2 - x_3 + 6x_4 + x_5 \geqslant 2 \\ &-3 \leqslant x_i \leqslant 3 \qquad \text{for } i = 1, 2, \ldots, 5 \ . \end{aligned}$$

[7] The points returned by the interior point algorithm are always feasible by construction; so we do *not* have to check explicitly where the constraint violation is small enough.

[8] As the transfer function is stable by construction, the stability requirement does not lead to any additional constraints on v.

So we have a nonlinear non-convex optimization problem with linear inequality constraints. The gradient and the Hessian of the objective function can be computed easily.
A suitable method is *multi-start* gradient projection with quadratic line minimization (M2) or — alternatively — a *multi-start* penalty function method with steepest descent line minimization (M9).

(b) The corresponding stopping criterion is: there exists a vector μ_k such that

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

where x_k is the current iteration point, f is the objective of the minimization problem, g is the inequality constraint function (when the constraints are written in the form $g(x) \leqslant 0$), and ε_1, ε_2, ε_3, ε_4 are small positive numbers.
For the penalty function approach we could also use $\|\nabla f_{\mathrm{tot}}(x_k)\|_2 \leqslant \varepsilon$ with $\varepsilon > 0$ and where f_{tot} is the sum of the original objective function and the penalty function.

P10. (a) Since $f_1 : x \mapsto \sqrt[3]{x}$ is a non-decreasing function, we can also minimize its argument:

$$\min_{x \in \mathbb{R}^4} \, (x_1 + 6x_2 + 8x_3 - 9x_4 - 10)^2 \ . \tag{2}$$

Next, we can use the constraint to eliminate x_1:

$$x_1 = \frac{5 - (3x_2^2 + 2x_2^2 x_3^2 + 4x_3^2 x_4^2 + 8x_4^4)}{1 + x_3^2} \ .$$

If we fill this out in the simplified objective function (cf. (2)), we get an unconstrained optimization problem with a non-convex nonlinear objective function. The gradient and the Hessian can be computed easily. Hence, the most appropriate algorithm is a *multi-start* line search method with the Levenberg-Marquardt direction.

(b) The corresponding stopping criterion is $\|\nabla f_{\mathrm{elim}}(x_k)\|_2 \leqslant \varepsilon$ where x_k is the current iteration point, f_{elim} is the objective function obtained by elimination of x_1, and $\varepsilon > 0$.

QUESTION 2: Optimization methods II

Answers

1. The following elements should be present in your answer:

   - The steepest descent direction in a point x_k is given by $-\nabla f(x_k)$ and it indicates the direction in which the function will locally decrease most rapidly in the neighborhood of x_k.
   - The general n-dimensional optimization problem is turned into a 1-dimensional optimization problem of the form

     $$s^* = \arg\min_s f(x_k - s\nabla f(x_k)) \quad .$$

     The new iteration point is then $x_{k+1} = x_k - s^* \nabla f(x_k)$.
     Optionally, it could be stated that the algorithm can be stopped if $\|\nabla f(x_k)\|_2 \leqslant \varepsilon$ for some small positive scalar ε.
   - A 2-dimensional[9] picture should be included with contour lines; this picture should illustrate that the gradient is orthogonal to the contour lines, the optimal point on a search line is found as the intersection with a contour for which the given search line is a tangent line, and as a consequence, all search lines are orthogonal to each other.
     In addition, optionally, a plot could be included that shows how the general n-dimensional optimization problem is turned into a 1-dimensional optimization problem.

2. If we denote the objective function by f, we have

   $$\nabla f = \begin{bmatrix} 4x - y - 8 \\ 2y - x \end{bmatrix} \quad \text{and} \quad H_f = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix} \quad .$$

   a) The objective function is convex since H_f is positive definite (as its leading principal minors are all positive) or since $f(x,y)$ can also be written as $\frac{7}{4}x^2 + \left(\frac{x}{2} - y\right)^2 - 8x + 10$, which is a sum of convex functions.

   b) The points where a local optimum is reached, are found by determining the points for which $\nabla f(x,y) = 0$. This yields

   $$\begin{cases} 4x - y - 8 = 0 \\ 2y - x = 0 \quad . \end{cases}$$

   From the second equation it follows that $x = 2y$; filling this out in the first equation yields $7y - 8 = 0$, or $y = \frac{8}{7}$.
   We thus find one optimum: $(x^*, y^*) = \left(\frac{16}{7}, \frac{8}{7}\right)$.
   The corresponding function value is $f(x^*, y^*) = \frac{6}{7}$.
   Since the function is convex, the point (x^*, y^*) is a (global) minimum.

[9] Note that if you only consider a function of only 1 variable, the line minimization actually is equivalent to the original optimization problem.

c) In the point $(x_0, y_0) = (0,0)$ we have $\nabla f(0,0) = \begin{bmatrix} -8 & 0 \end{bmatrix}^{\mathrm{T}}$. Hence, we have to search along the line defined by $x = 0 + 8s$, $y = 0$, or equivalently $x = t$, $y = 0$. This results in

$$\min_t \; 2t^2 - 8t + 10 \;\; .$$

The optimal t value is found by setting the derivative equal to 0, or $4t - 8 = 0$, which yields $t^* = 2$. This results in the following new iteration point: $(x_1, y_1) = (2,0)$.
We have $f(2,0) = 2$.

Since $\nabla f(2,0) = \begin{bmatrix} 0 & -2 \end{bmatrix}^{\mathrm{T}}$, we next have to search along the line defined by $x = 2$, $y = 2s$, or equivalently $x = 2$, $y = t$. This results in

$$\min_t \; 2 \cdot 4 + t^2 - 2t - 16 + 10 \;\; = \;\; \min_t \; t^2 - 2t + 2 \;\; .$$

The optimal t value is found by setting the derivative equal to 0: $2t - 2 = 0$, which yields $t^* = 1$. This results in the following new iteration point: $(x_2, y_2) = (2,1)$.
We have $f(2,1) = 1$.

QUESTION 3: Controller design

Answers

a) We have[10]

$$e = r - (P_2Ke + P_1(d + Ke)) = r - P_1d - P_2Ke - P_1Ke \ ,$$

which results in

$$e = \frac{1}{1 + (P_1 + P_2)K}(r - P_1d) \ .$$

This leads to

$$u = Ke = \frac{1}{1 + (P_1 + P_2)K}(Kr - KP_1d)$$

and

$$\begin{aligned} y = P_1(d + u) &= P_1d + \frac{P_1}{1 + (P_1 + P_2)K}(Kr - KP_1d) \\ &= \frac{1}{1 + (P_1 + P_2)K}\left(KP_1r + \left(P_1 + (P_1 + P_2)KP_1 - KP_1^2\right)d\right) \\ &= \frac{1}{1 + (P_1 + P_2)K}(KP_1r + P_1(1 + P_2K)d) \ . \end{aligned}$$

Hence,

$$M = \frac{1}{1 + (P_1 + P_2)K}\begin{bmatrix} K & -KP_1 \\ KP_1 & P_1(1 + P_2K) \\ 1 & -P_1 \end{bmatrix} \ .$$

b) The closed-loop system will be internally stable if all transfer functions from any external input (in our case r and d) to any internal signal are stable. Apart from the signals e, y, and u, the other internal signals in the given system are the output of P_2, and the outputs of the middle and the rightmost summing junctions in the figure. However, since P_2 is stable, and since addition preserves stability, it is sufficient to show that any controller in $\mathscr{K}$ will result in stable transfer matrix M.
First we substitute the given expression for K into M. Since

$$\begin{aligned} \frac{1}{1 + (P_1 + P_2)K} &= \frac{1}{1 + (P_1 + P_2)\frac{Q}{1-(P_1+P_2)Q}} \\ &= \frac{1 - (P_1 + P_2)Q}{1 - (P_1 + P_2)Q + (P_1 + P_2)Q} \\ &= 1 - (P_1 + P_2)Q \end{aligned}$$

[10]For the sake of brevity of notation we drop the arguments k and q.

and

$$
\begin{aligned}
P_1(1+P_2K) &= P_1\left(1+P_2\frac{Q}{1-(P_1+P_2)Q}\right) \\
&= P_1\frac{1-(P_1+P_2)Q+P_2Q}{1-(P_1+P_2)Q} \\
&= \frac{P_1(1-P_1Q)}{1-(P_1+P_2)Q} \ ,
\end{aligned}
$$

we have

$$
M = \begin{bmatrix} Q & -QP_1 \\ QP_1 & P_1(1-P_1Q) \\ 1-(P_1+P_2)Q & -P_1(1-(P_1+P_2)Q) \end{bmatrix} \ .
$$

Since M contains sums and products of P_1, P_2, and Q, which are all stable transfer functions, and since stability is preserved under addition and multiplication, all entries of M are also stable transfer functions.
So any controller in $\mathscr{K}$ indeed internally stabilizes the closed-loop system.