Worked solutions for Sample Exam 1
“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

P1. (a) Although the objective function of this problem is convex, the constraint $x_1^2 + x_2^2 \geqslant 1$ is not convex, since in the standard form we get $-x_1^2 - x_2^2 + 1 \leqslant 0$, which has a concave (i.e., the opposite of convex) left-hand side. 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 (M9).

(b) The most appropriate stopping criterion is[1]: there exists a 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, and g is the inequality constraint function (written in the form $g(x) \leqslant 0$).

P2. (a) By introducing a dummy variable t, the minimization problem $\min_x |x_1 - 4x_2 + x_3 - 4|$ can be recast into the constrained problem $\min_{t,x} t$ subject to $t \geqslant |x_1 - 4x_2 + x_3 - 4|$ or equivalently $t \geqslant x_1 - 4x_2 + x_3 - 4$ and $t \geqslant -(x_1 - 4x_2 + x_3 - 4)$, which are linear constraints. Note that the constraints $x \geqslant 0$ and $x_1 + x_2 \geqslant 25$ are also linear. So we get a linear programming problem and therefore the most 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[2].

[1]There are no equality constraints; so the equality constraint function h does not appear here.
[2]However, in practice a maximum number of iterations is usually specified.

P3. (a) The objective function contains an integral that is most probably not expressible in closed form and that should thus be computed numerically. Moreover, the objective function is not convex (due to the non-convex terms in the argument of the integral). So we have a nonlinear non-convex optimization problem with nonlinear (non-convex) constraints. Therefore, a multi-start optimization method is required. The gradient and Hessian of the objective function cannot be computed analytically. The numerical computation of the integral will in general be time-consuming and as such it is better not to use numerical computation of the gradient and Hessian of the objective function, but rather select a gradient-free optimization method. So in this case a *multi-start* penalty function method with the Nelder-Mead algorithm (M7) is the most suited optimization algorithm.

(b) A suitable stopping criterion is[3]

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

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

P4. (a) The objective function contains the non-convex term $x_1^2 x_2^2$. Note that this term is not convex since it is 0 on the x_1-axis and on the x_2-axis and strictly positive elsewhere; so, e.g., the line connecting the points $(1,0)$ and $(0,1)$ does not lie above the function surface. In principle[4] we thus have a nonlinear, non-convex optimization problem. This implies that a multi-start method is required. The gradient and the Hessian of the objective function can easily be computed analytically. So in this case the *multi-start* Levenberg-Marquardt algorithm (M5) is the most suited optimization algorithm.

(b) The most appropriate stopping criterion is

$$\|\nabla f(x_k)\|_2 \leqslant \varepsilon \ ,$$

where x_k is the current iteration point and f is the objective function.

P5. (a) The objective function of this optimization problem is convex in its argument x since it is the sum of two convex functions:

* $\cosh(x_1^2 + x_2^2)$ is convex since $x_1^2 + x_2^2$ is convex (and nonnegative), since the function h defined by $h(y) = \cosh y$ is convex and non-decreasing for $y \geqslant 0$, and since a composed function $h(g(x)$ is convex in its argument x if g is convex and h is convex and non-decreasing;
* $14x_2^4$ is convex.

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

[4] In this particular case, it turns out that the function is convex since the Hessian is positive definite. However, this requires extensive symbolic and numerical computations, and so we do not expect you to discover this. As a result, in fact the best suited method is the cutting-plane algorithm (M3) with as stopping criterion is $|f(x_k) - f(x^*)| \leqslant \varepsilon_f$, where x_k is the current iteration point, x^* is the (yet unknown) optimum of this optimization problem, and f denotes the objective function.

The constraint is also convex since $x_1^2 + (x_2 - 10)^2 - 50$ is the sum of several convex functions. Hence, the optimization problem is convex. This implies that a suited optimization algorithm is the cutting-plane algorithm (M3).

(b) For the cutting-plane method a suited stopping criterion is

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

where x_k is the current iteration point, x^* is the (yet unknown) optimum of this optimization problem, f denotes the objective function, and g denotes the constraint function g (with the constraint written in form $g(x) \leqslant 0$).

P6. (a) This problem is very similar to the one of P5: only the direction of the inequality sign in the constraint has changed. As a result we now have a non-convex constraint, and thus a nonlinear, non-convex constrained optimization problem. This implies that a multi-start method is required. The gradient and the Hessian of the objective function can easily be computed analytically. So in this case it is best to select *multi-start* sequential quadratic programming (M9).

(b) The most appropriate stopping criterion is: there exists a 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, and g is the inequality constraint function (written in the form $g(x) \leqslant 0$).

P7. (a) The objective function is non-convex due to the term $-x_2^3$. So we have a nonlinear, non-convex constrained optimization problem. Therefore, a multi-start method is required. The gradient and the Hessian of the objective function can easily be computed analytically. So in this case it is best[5] to select *multi-start* sequential quadratic programming (M9).

(b) The most appropriate stopping criterion is:there exists a 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, and g is the inequality constraint function (written in the form $g(x) \leqslant 0$).

[5]Note that sequential quadratic programming (M9) uses 2nd-order information (the Hessian), while the gradient projection method with Fibonacci line search (M8) only uses 1st-order information (the gradient). Therefore, M9 is preferred to M8.

P8. (a) Since the function g defined by $g(y) = 2^y$ is non-decreasing, the problem can be simplified to the minimization of $5x_1^2 + 2x_2^2 + 6x_1x_2 + 3x_1 - x_2$, which can also be rewritten as $(2x_1 + x_2)^2 + (x_1 + x_2)^2 + 3x_1 - x_2$ and which is thus[6] a convex quadratic function. Since the constraints are linear, we thus have to solve a quadratic programming 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[7].

P9. (a) The objective function of this optimization problem is convex in its argument x since it is the sum of two convex functions:

* $((x_1 - 1)^2 + (x_2 - 2)^2)^2$ is convex since $(x_1 - 1)^2 + (x_2 - 2)^2$ is convex (and non-negative), since the function h defined by $h(y) = y^2$ is convex and non-decreasing for $y \geq 0$, and since a composed function $h(g(x)$ is convex in its argument x if g is convex and h is convex and non-decreasing;
* $-\log(2 - x_1^2 - x_2^2)$ is convex since the function g defined by $g(x) = 2 - x_1^2 - x_2^2$ is concave (i.e., $-g$ is convex), since the function h defined by $h(y) = -\log(y)$ is convex and non-increasing, and since a composed function $h(g(x)$ is convex in its argument x if $-g$ is convex and h is convex and non-increasing.

The constraint is also convex since $\|x\|_1$ is a norm function and norm functions are convex in their argument. Hence, this optimization problem is convex. So the best suited optimization algorithm is the cutting-plane algorithm (M3).

(b) For the cutting-plane method a suited stopping criterion is

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

where x_k is the current iteration point, x^* is the (yet unknown) optimum of this optimization problem, f denotes the objective function, and g denotes the constraint function g (with the constraint written in form $g(x) \leqslant 0$).

P10. (a) The objective function is non-convex due to the term $(x_1 + x_2)^3$. So we have a non-linear, non-convex constrained optimization problem. The gradient and the Hessian of the objective function can easily be computed analytically. So in this case it is best to select *multi-start* sequential quadratic programming (M9).

[6] An alternative way to show that the quadratic function $5x_1^2 + 2x_2^2 + 6x_1x_2 + 3x_1 - x_2$ is convex, is to determine its Hessian $H = \begin{bmatrix} 10 & 6 \\ 6 & 4 \end{bmatrix}$, which is positive definite since its leading principal minors (10 and $10 \cdot 4 - 6 \cdot 6 = 4$) are both positive.

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

(b) The most appropriate stopping criterion is:there exists a 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, and g is the inequality constraint function (written in the form $g(x) \leqslant 0$).

QUESTION 2: Optimization methods II

Answer for Question a)

- The absolute value function $|.|$ is convex (this can be verified by plotting it or better by noting it is a norm function). A convex function of a linear argument is also convex and the sum of two convex functions is also convex. Hence, f is a convex function.

- Note that the function f always has a nonnegative value. The first term $|2x-2y|$ equals 0 if $x=y$ and the second term $|x+y-2|$ equals 0 if $x+y=2$. So $f(x,y)$ is equal to 0 if and only if $x=y$ and $x+y=2$, or equivalently $(x,y)=(1,1)$. For any other point either the first term or the second term will be larger than 0, and therefore $(1,1)$ is the unique global minimum of f.

- If we optimize along the first coordinate direction starting from the point $(0,0)$ we have to replace (x,y) by $(0,0)+s(1,0)=(s,0)$ with s the step size, and optimize $F(s)=f(s,0)=|2s|+|s-2|$. This function can be rewritten is

$$F(s)=\begin{cases} -2s-s+2=2-3s & \text{if } s\leqslant 0 \\ 2s-s+2=s+2 & \text{if } s\geqslant 0 \text{ and } s\leqslant 2 \\ 2s+s-2=3s-2 & \text{if } s\geqslant 2\ . \end{cases}$$

  Hence, the minimum of F is reached for $s^*=0$ (with $F(0)=2$). So we stay in the initial point $(0,0)$.
  If we optimize along the second coordinate direction starting from the point $(0,0)$, we have to replace (x,y) by $(0,0)+s(0,1)=(0,s)$ with s the step size, and optimize $G(s)=f(0,s)=|-2s|+|s-2|$. This function is in fact identical to the function F considered in the previous line minimization. Hence, the minimum of G is also reached for $s^*=0$. So we still stay in the initial point $(0,0)$ after one cycle of the perpendicular search method, which implies that the method fails.

Answer for Question b)

- The condition
$$\forall y\in[0,1]:\begin{bmatrix} \gamma-y-1 & x \\ x & y+1 \end{bmatrix}\geqslant 0$$
  is a linear matrix inequality (LMI) in y, and since LMIs are convex, we know that if the condition holds for the end points $y=0$ and $y=1$ it will also hold for all intermediate points $y\in(0,1)$. So the condition is satisfied if and only if

$$\begin{bmatrix} \gamma-1 & x \\ x & 1 \end{bmatrix}\geqslant 0$$
$$\begin{bmatrix} \gamma-2 & x \\ x & 2 \end{bmatrix}\geqslant 0\ .$$

- First we use the result of step 1 to recast the problem into

$$\begin{aligned}
&\min_{\gamma, x \in \mathbb{R}} \gamma \\
&\forall y \in [0,1] : \begin{bmatrix} \gamma - y - 1 & x \\ x & y+1 \end{bmatrix} \geqslant 0 \ .
\end{aligned}$$

Since for $y \in [0,1]$ the term $y+1$ is always invertible, we can now apply the Schur complement property to

$$\begin{bmatrix} \gamma - y - 1 & x \\ x & y+1 \end{bmatrix} \geqslant 0 \ ,$$

which yields the equivalent conditions

$$\begin{aligned}
&y+1 \geqslant 0 \\
&\gamma - y - 1 - \frac{x^2}{y+1} \geqslant 0
\end{aligned}$$

or, since $y+1 \geqslant 0$ is always satisfied for $y \in [0,1]$, just

$$\gamma \geqslant y + 1 + \frac{x^2}{y+1} \ .$$

Since this condition should hold for all $y \in [0,1]$, we have

$$\gamma \geqslant \max_{y \in [0,1]} f(x,y)$$

So now we have the problem

$$\begin{aligned}
&\min_{x, \gamma \in \mathbb{R}} \gamma \\
&\gamma \geqslant \max_{y \in [0,1]} f(x,y) \ .
\end{aligned}$$

Since γ will be minimized and since γ only appears in the left-hand side of the constraint, this implies that the optimal γ will be equal to the right-hand side of the constraint. Hence, the given optimization problem is equivalent to

$$\min_{x \in \mathbb{R}} \max_{y \in [0,1]} y + 1 + \frac{x^2}{y+1} \ .$$

QUESTION 3: Controller design

Answer for Question a)

For the given input signal we have the following difference equation:

$$y(k) - y(k-1) = a$$

with $y(0) = y_0$. To solve this equation we first compute the homogeneous solution y_{hom}, i.e., a solution of $y(k) - y(k-1) = 0$. This gives $y_{\text{hom}}(k) = 1$ for all k. Next, we compute a particular solution y_{part}. Since we have $y(k) = a + y(k-1)$, it is easy to verify that $y_{\text{part}}(k) = ka$ for all k is particular solution. The general solution is thus $y(k) = Cy_{\text{hom}}(k) + y_{\text{part}}(k) = C + ka$ with C a constant. The initial condition $y(0) = y_0$ yields $C = y_0$, and thus we find

$$y(k) = y_0 + ka \quad \text{ for } k = 0, 1, 2, \ldots$$

For a fixed k, this function is affine in a and thus we conclude that $y(k)$ is convex as a function of a.

We have

$$y_{\text{av}} = \frac{1}{N} \sum_{k=1}^{N} (y_0 + ka) = \frac{1}{N} \left(Ny_0 + a\frac{N(N+1)}{2} \right) = y_0 + \frac{N+1}{2} a \ .$$

This expression is affine in a and thus we conclude that y_{av} is convex as a function of a.

Answer for Question b)

Now we have

$$y(k) - by(k-1) = 1 \ .$$

The homogeneous solution is now given by $y_{\text{hom}}(k) = b^k$. As particular solution we find

$$y_{\text{part}}(k) = \begin{cases} k & \text{if } b = 1 \\ \frac{1}{1-b} & \text{otherwise.} \end{cases}$$

So we get

$$y(k) = \begin{cases} y_0 + k & \text{if } b = 1 \\ \left(y_0 - \frac{1}{1-b}\right) b^k + \frac{1}{1-b} & \text{otherwise.} \end{cases}$$

Clearly, $y(k)$ is not convex as a function of b due to the appearance of the non-convex term $\frac{1}{1-b}$. Likewise, y_{av} is not convex as a function of b.

Answer for Question c)

By selecting $c = 0$ and $d = 1$, we retrieve the previous case, and since there we showed that $y(k)$ and y_{av} are not convex as a function of b, we can conclude that in this case $y(k)$ and y_{av} are not convex as a function of b, c, and d.