Exam — October 2021 – Grading template
Optimization for Systems and Control (SC42056)

Important: Please recall the following instructions from the exam procedure:

- Note that – just as in previous years – **correct results without proper and correct motivation will not receive any marks.**

For an example on how proper and correct motivations look like, please consult the worked solutions for Sample Exams 1 and 2 and for the exams of October 2013 and October 2014

Additional scoring guidelines

S0: correct result without proper and correct motivation: 0

S1a: small computation error that *does* affect result: -0.5

S1b: small computation error that does *not* affect result: -0.25

S2: partially incomplete motivation for convexity or simplification: -50%

S3: missing, wrong, or not properly motivated $N\times$ in Question 1: score 0 for (c) and penalty -1 for (b)

S4: multi-start listed when it is not needed: -0.5

S5: ∇f as row vector: -0.5

S6: if in Question 1 the answer for (c) and/or (d) is formally correct, but an error is made in (a) or (b) that affects the result for (c) and/or (d): 0 for (c) and/or (d)

S7: redundant function in stopping criterion: 0

S8: introduction of redundant variables or constraints that are not needed at all: -0.75

S9: introduction of additional wrong constraints and/or wrong classification of extra/unsimplified constraint: -1

CE: Even if the answers to (a)-(d) are wrong, you can still score marks for (e) if and only if (a)-(d) are internally consistent and all result in the answer given in (e) and if (e) is 100 % correct and complete.

On the next pages **concise** answers are given with scores marked in red. To earn the indicated score the corresponding answer has to be given completely, including the information inside the brackets; else the score is 0.

QUESTION 1 ($9 \times 8 = 72$ points)

- **P1**

1 (a) The objective function should not be simplified : 0.5
The first constraint can be rewritten as: $\|x\|_2 \leq \sqrt[4]{1000}$ (1)
As $(\cdot)^5$ is a nondecreasing function, the second constraint can be simplified to
$x_1^6 + x_2^2 - 3x_3 + 8x_4^2 - x_5x_7 + 4x_5^2 + 8x_7^2 + x_9 - x_{10} \leq 2$ (2)
0.5 (for both (1) and (2) together, with penalty of -0.25 if it is not mentioned that $(\cdot)^5$ is nondecreasing when simplifying the second constraint)

4 (b) The function $A_1(3\cdot)$, $2A_2(\cdot)$, $A_3(6\cdot) + 8A_4(2\cdot)$ are convex, as a convex function of a linear argument is convex and as the positively weighted sum of convex functions is convex
Since $3^{(\cdot)}$ is convex and increasing, and since its argument is convex, the function $3^{A_3(6\cdot)+8A_4(2\cdot)}$ is also convex
Since the objective function is the maximum of convex functions is also convex: 2
Constraint (1) as $\|\cdot\|_2$ is convex: 0.5
If constraint (1) was not simplified yet, it should mentioned that the constraint is convex as $(\cdot)^4$ is convex and increasing (for nonnegative arguments) and $\|\cdot\|_2$ is convex.
Constraint (2) is convex as is the sum of even powers, a square of a linear term (i.e., $(x_5 - 0.5x_7)^2$), and an affine term, all of which are convex : 1.5

1 (c) CP: convex optimization problem

1.5 (d) We have a convex problem, but the gradient of the objective function cannot be computed analytically; moreover, due to the high number of variables and the long time it takes to evaluate the objective function, numerical computation of the gradient would be very time-consuming; hence, a gradient-free method is recommended. So the best method is:
M10: Penalty function approach + Powell : 1; **no** multi-start needed as Powell can follow the shape of the function : 0.5 **or**
multi-run : 0.5 + M11: Simulated annealing : 1

0.5 (e) $|f_s(x_k) - f_s(x_{k-1})| \leq \varepsilon_f$ and $g_s(x_k) \leq \varepsilon_g$ (for M10: Penalty function + Powell) **or**
$T_k < T_{\text{final}}$ (for M11: Simulated annealing)

- **P2**

1 (a) The objective function cannot be simplified : 0.25
Constraint (1), (3), and (4) can stay as they are : 0.25
Actually, (2) makes (3) redundant.
As $\log(\cdot)$ is an increasing function (and its argument its positive due to constraint (3)), the second constraint can be rewritten as $20-8x_1^2-6x_2^2-x_3-x_4+2x_5 \geq e^9$ or equivalently $8x_1^2+6x_2^2+x_3+x_4-2x_5 \leq 20-e^9$ (2) : 0.5

4 (b) As the argument of the $\cosh(\cdot)$ is not always nonnegative, we cannot show that the first term of the objective function is convex; actually, it is even nonconvex (this can easily be shown by considering the simplified case $\cosh(x^2-y)$ (so $x_1=x$, $x_2=x_3=0$, $2x_5=-y$): we get a function value $\cosh(0)=1$ for the points $(1,1)$ and $(-1,1)$, while the intermediate point $(0,1)$ has a function value $\cosh(-1)\geq 1$. So the line connecting 2 points on the function surface is not below the function surface: 3
Constraint (1) is linear
Constraint (2) and (3) are similar and they are both convex as the positive sum of squares and a linear function is convex : 0.5
Constraint (4) is convex as $(\cdot)^4$ is convex and increasing for nonnegative arguments, and as the function $(x_1,x_2)\mapsto x_1^2+x_2^2$ is convex and nonnegative, and as the sum of a convex function and an affine function is convex: 0.5

1 (c) NCC: nonconvex constrained optimization problem

1.5 (d) as the gradient (subgradient for the second term (which is convex) of the objective function) and the hessian can be computed analytically, the best choice is:
multi-start: 0.5 + M9: Barrier function approach + steepest descent

0.5 (e) KKT conditions with ε (list them!) **<u>or</u>**
$\|\nabla f_{\text{barrier+}}(x_k)\|_2 \leq \varepsilon$ (where $f_{\text{barrier+}}$ is the sum of the objective function f and the barrier function)

- **P3**

1 (a) As the argument is always positive, we can get rid of the exponent -5 (as it is then a decreasing function) and of the constant 7, resulting in a **maximization** problem:
$\mathbf{max}_{x\in\mathbb{Z}^5} 3|x_1|+2|x_2|+8|x_3|+6|x_4|+3|x_5|$ or equivalently
$\mathbf{min}_{x\in\mathbb{Z}^5} -3|x_1|-2|x_2|-8|x_3|-6|x_4|-3|x_5|$: 1
As the weights in the objective function are negative we **cannot** use the trick with dummy variables to obtain a linear objective function; if that is done anyway the score for (a) is 0.
for simplifications of the constraints, see (b)

4 (b) We can either consider $2^5 = 32$ different problems with linear objective functions $\pm 3x_1 \pm 2x_2 \pm 8x_3 \pm 6x_4 \pm 3x_5$ **or** — taking into account the analysis of the first constraint and in view of computational efficiency (see also item (c)), — consider the objective function to be nonconvex (due to the negative coefficients in front of the convex terms $|x_i|$): 1.5
Constraint (1) is nonconvex (as it can be rewritten as $4 - \|x\|_\infty \leq 0$, i.e., with a minus sign in front of the convex terms $\|x\|_\infty$), *but* can be written as the **union** of $2 \cdot 5 = 10$ linear constraints: $4 \leq x_1$ OR $4 \leq -x_1$ OR $4 \leq x_2$ OR ... OR $4 \leq -x_5$: 1.75
Constraint (2) can be written as $3 + 2x_1 + x_2 - 5x_3 + 9x_4 - 2x_5 \geq 10$, which is a linear constraint: 0.25
Constraint (3) can be written as a set (intersection!) of $2^5 = 32$ linear constraints $\pm 5x_1 \pm 4x_2 \pm 3x_3 \pm 3x_4 \pm x_5 \leq 50$: 0.5

1 (c) 32 (due to objective function) $\times$ 10 (due to constraint (1)) = 320 $\times$ MILP: mixed-integer linear programming problem **or**
NCC: nonconvex constrained optimization problem
if MILP is specified, but the objective function was not identified or not correctly proven (as indicated above) to be 32 $\times$ linear in (b), then S6 applies and the scores for (c) and (d) will be 0

1.5 (d) We need to select an algorithm that can deal with integer variables. So the most appropriate algorithm is:
M12: Branch-and-bound method for mixed-integer linear programming **or**
multi-run: 0.5 M11: Simulated annealing: 1

0.5 (e) optimum found once entire tree is explored (for M11) **or**
$T_k < T_{\text{final}}$ (for M11: Simulated annealing)

- **P4**

1 (a) Since $\sqrt[7]{(\cdot)}$ is a nondecreasing function, we can also minimize its argument:

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

Note that the square has to stay! 1
We keep the constraint as is.

4 (b) The objective function f_s is convex quadratic: 2
The constraint is nonconvex as it is an equality constraint and the left-hand side is not an affine function: 2 S0 applies the latter is not stated explicitly

1 (c) NCC: nonconvex constrained optimization problem

1.5 (d) As the gradient and the Hessian can be computed analytically, the best choice is: multi-start: 0.5 + M8: Lagrange method + Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm: 1

0.5 (e) $\|\nabla f_s(x_k) + \nabla h_s(x_k)\lambda\| \le \varepsilon_1$ and $\|h_s(x_k)\| \le \varepsilon_2$

P4 (alternative solution, with elimination)
This solution has 1 variable less, but it uses a first-order method (steepest descent) instead of a second-order one (BFGS)

1 (a) Since $\sqrt[7]{(\cdot)}$ is a nondecreasing function, we can also minimize its argument:

$$\min_{x\in\mathbb{R}^4} f_1 := \min_{x\in\mathbb{R}^4} (x_1 + 6x_2 + 8x_3 - 9x_4 - 10)^2$$

Note that the square has to stay! 0.25
Next, we can use the constraint, which is a quadratic function in x_4, to eliminate x_4:

$$x_4 = \pm\sqrt{\frac{1}{8}\left(100 - \exp(2x_1^2 + 2x_2^2 + 2x_1^2x_2^2 + x_3^4) - 2x_1^2 - x_2^4 - x_3^6\right)} \quad 0.5$$

If we fill out each of the 2 solutions into the quadratic objective function f_1 defined above and if we add the constraint that the argument of the square root has to be nonnegative: $100 - \exp(2x_1^2 + 2x_2^2 + 2x_1^2x_2^2 + x_3^4) - 2x_1^2 - x_2^4 - x_3^6 \ge 0$ or equivalently

$$\exp(2x_1^2 + 2x_2^2 + 2x_1^2x_2^2 + x_3^4) + 2x_1^2 + x_2^4 + x_3^6 - 100 \le 0 \qquad (1)$$

we get two problems of the form $\min_{x\in\mathbb{R}^3} f_s(k)$ s.t. (1) 0.25

4 (b) The objective function f_s defined above is nonconvex since x_4 is a nonconvex function of x_1, x_2, and x_3 : 2
Constraint (1) of item (a) is not convex as the product $x_1^2x_2^2$ cannot be absorbed in a sum of squares: 2

1 (c) $2\times$ NCC nonconvex constrained optimization problem
if only one solution for x_4 was provided in (a), then S3 applies, and the score will be 0 for (c) and a penalty -1 holds for (d)

1.5 (d) as the gradient and the Hessian can be computed analytically, the best choice is: multi-start: 0.5 + M9: Barrier function approach + steepest descent: 1

0.5 (e) KKT conditions with ε (list them!) <u>**or**</u>
$\|\nabla f_{\text{barrier+}}(x_k)\|_2 \le \varepsilon$ (where $f_{\text{barrier+}}$ is the sum of the objective function f_s and the barrier function)

- **P5**

1 (a) As $\sinh(\cdot)$ is an increasing function we can maximize its argument; next, the constant 1 can be dropped: 0.5
as $-\cosh$ is a decreasing function for **nonnegative** arguments (note that the term x_3x_4 can be absorbed), we can minimize the argument. So we finally get $\min_x f_s(x) := \min_{x\in\mathbb{R}^4} 2x_1^2 + 4x_2^2 + x_3^2 + (x_3 - 4x_4)^2$: 0.5
for simplifications of the constraints, see (b)

4 (b) The objective function f_s can be written as a sum of squares (see (a)), so it is a convex quadratic function: 2
Constraint (1) can be written as the **union** of 2 linear constraints:
$-25 \leq x_1 + 2x_2 + 3x_3 + 4x_4 \leq -1$ OR $1 \leq x_1 + 2x_2 + 3x_3 + 4x_4 \leq 25$: 1
Since an even power is a U-shaped function, constraint (2) can be transformed into the **union** of 2 linear constraints: $7x_1 + \ldots \leq -2$ OR $7x_1 + \ldots \geq 2$: 1

1 (c) $4\times$ QP: convex quadratic programming problem

1.5 (d) M3: Interior point algorithm: 1
as there is no dedicated QP algorithm: 0.5

0.5 (e) $|f_s(x_k) - f_s(x^*)| \leq \varepsilon_f$ **and** $g_s(x_k) \leq \varepsilon_g$
Actually, for the interior point method the constraint is always satisfied; but you have to say this explicitly.

- **P6**

1 (a) By introducing $x_1 = \cos\theta_1$, $x_2 = \sin\theta_2$, $x_3 = \theta_3$, $x_4 = \sin\theta_4$, $x_5 = \theta_5$, $x_6 = \sin\theta_6$, we get the following problem:

$$\min_{x\in\mathbb{R}^6} \sum_{k=1}^{7500} \varepsilon^2(k,x) \qquad \text{s.t. } x_i \in [0,1] \text{ for } i = 1,2,4,6 \text{ and } x_j \in [0,\pi/2] \text{ for } j = 3,5$$

and with a discrete-time model

$$y(k) = x_1 \cdot e^{2y(k-1)} + x_2 \cdot y(k-3)\cos(u(k)) + x_3|u(k-3)| + \\ x_4 u^4(k-2) + x_5 y(k-4)\sin(y(k-1)) + x_6 \cdot \frac{u^2(k-1)}{1+y^2(k-2)u(k-2)}$$

4 (b) note that $y(k)$ and thus also $\varepsilon(k,x)$ depend linearly on x, which means that the objective function is convex quadratic: 3.5
the constraints on x_i are simple bound constraints, and thus they linear: 0.5

1 (c) QP: convex quadratic programming problem

1.5 (d) M3: Interior point algorithm: 1
as there is no dedicated QP algorithm: 0.5

0.5 (e) $|f_s(x_k) - f_s(x^*)| \leq \varepsilon_f$ **and** $g_s(x_k) \leq \varepsilon_g$
Actually, for the interior point method the constraint is always satisfied; but you have to say this explicitly.

- **P7**

1 (a) Maximizing $2-\arctan(\cosh(\dots))$ is equivalent to minimizing $\arctan(\cosh(\dots)$. Since $\arctan(\cdot)$ is a nondecreasing function, we can also minimize $\cosh(\dots)$ instead: 0.5
The argument of the $\cosh(\dots)$ part can be rewritten as $f_s(x)=\frac{11}{4}x_1^2+(\frac{1}{2}x_1+x_2)^2+(2x_3-2x_4)^2+x_4^2$, which is nonnegative. Since $\cosh(\cdot)$ is increasing for nonnegative arguments, we can minimize its argument instead: 0.5
for simplifications of the constraints, see (b)

4 (b) as the objective function f_s is a positive sum of squares of linear functions, it is a convex quadratic function: 1.5
Since $\exp(\cdot)$ is nondecreasing and since $\exp(a)\exp(b)=\exp(a+b)$, the first constraint can be rewritten as $3+2x_1+4x_2+x_3-2x_4\leq 17$, which is linear : 0.5
The second constraint can be written as a set (intersection!) of 2 linear constraints: $2x_1+x_2+3x_3+x_4\geq 12$ **and** $7x_1-x_2+6x_3-6x_4\geq 12$: 1
The third constraint consists of a positive sum of even power (which are convex) and therefore it is a convex constraint: 1

1 (c) CP: convex optimization problem

1.5 (d) M3: Interior point algorithm

0.5 (e) $|f_s(x_k)-f_s(x^*)|\leq\varepsilon_f$ **and** $g_s(x_k)\leq\varepsilon_g$
Actually, for the interior point method the constraint is always satisfied; but you have to say this explicitly.

- **P8**

  1 (a) no simplification possible

  4 (b) Due to presence of cos: 1.5, the objective function will be a nonconvex function: 1.5
  Note that sinh is convex for nonnegative arguments, so the presence of sinh alone cannot be used to motivate nonconvexity
  The constraints are simple bound constraints and thus linear : 1

  1 (c) NCC: nonconvex constrained optimization problem

  1.5 (d) The integral that appears in the objective function cannot be computed analytically, so numerical computation is required, which will be time-consuming. Therefore, and also due to the high number of variables, a gradient-free method is recommended. So the best choice is:
  multi-start: 0.5 + M10: Penalty function approach + Powell : 1 **or**
  multi-run : 0.5 + M11: Simulated annealing: 1

  0.5 (e) $|f_s(x_k) - f_s(x_{k-1})| \leq \varepsilon_f$ and $g_s(x_k) \leq \varepsilon_g$ (for M10: Penalty function + Powell) **or**
  $T_k < T_{\text{final}}$ (for M11: Simulated annealing)

- **P9**

1 (a) Since $(\cdot)^4$ is increasing for nonnegative arguments, and since the argument is indeed nonnegative as it is a sum of absolute values, we can rewrite the problem as $\min_{x\in\mathbb{R}^4} 4|x_1| + 3|x_2| + |x_3| + 2|x_4| + 1$: 0.5
The constant 1 can be dropped, and we can introduce dummy variables α_i, $i = 1,2,3,4$ with $\alpha_i \geq |x_i|$ or equivalently $\alpha_i \geq x_i$ and $\alpha_i \geq -x_i$ for $i = 1,2,3,4$. Then we get a problem of the form $\min_{x,\alpha} 4\alpha_1 + 3\alpha_2 + \alpha_3 + 2\alpha_4$ subject to the given constraint as well as $\alpha_i \geq x_i$, $\alpha_i \geq -x_i$ for $i = 1,2,3,4$: 0.5
for simplifications of the constraints, see (b)

4 (b) The objective function is linear (in α): 1
As $\exp(\cdot)$ is an increasing function we can rewrite the first constraint as $(6+3x_1+x_2-x_3-x_4)^4 \leq \log(16000)$. Since an even power is a U-shaped function, this constraint can be transformed into the intersection of 2 linear constraints: $-\sqrt[4]{\log(16000)} \leq 6+3x_1+x_2-x_3-x_4 \leq \sqrt[4]{\log(16000)}$: 0.5
Note that the 4th power applies to the expression in the brackets, not to exp; in the latter case we would have to write $\exp^4(\ldots)$ or $[\exp(\ldots)]^4$
The second constraint can be written as a set (intersection!) of 3 linear constraints: $3+2x_1+3x_2-x_4 \leq 9$ **and** $4x_2+3x_4 \leq 9$ **and** $2x_1 - 8 \leq 9$: 0.5
The third constraint can be written as $\max(|x_1|,|x_2|,|x_3|,|x_4|)+|x_1|+|x_2|+|x_3|+|x_4| \geq 2$. or $\max(2|x_1|+|x_2|+|x_3|+|x_4|, |x_1|+2|x_2|+|x_3|+|x_4|, |x_1|+|x_2|+2|x_3|+|x_4|, |x_1|+|x_2|+|x_3|+2|x_4|) \geq 2$, or as the **union** of 4 constraints: $2|x_1|+|x_2|+|x_3|+|x_4| \geq 2$ **or** $2|x_1|+2|x_2|+|x_3|+|x_4| \geq 2$ **or** $2|x_1|+|x_2|+2|x_3|+|x_4| \geq 2$ **or** $2|x_1|+|x_2|+|x_3|+2|x_4| \geq 2$. Each of these 4 constraints can in its turn be written as the **union** of $2^4 = 16$ linear constraints: e.g., for the first one we get $2x_1+x_2+x_3+x_4 \geq 2$ **or** $2x_1+x_2+x_3-x_4 \geq 2$ **or** $2x_1+x_2-x_3+x_4 \geq 2$ **or** … **or** $-2x_1-x_2-x_3-x_4 \geq 2$. So third constraint leads to a **union** of $4 \cdot 16 = 64$ linear constraints: 2
Note that we cannot substitute α_i in these constraints since these would result in additional lower bounds for α_i, as a consequence of which the relation $\alpha_i^* = |x_i^*|$ can no longer be guaranteed.

1 (c) $64 \times$ LP

1.5 (d) M1: Simplex algorithm for linear programming

0.5 (e) The simplex algorithm will always find a global optimum in a finite number of iterations

QUESTION 2 $(9+16+3=28$ points)

- **Question 2.1**

7 (a) Mention/provide at least the following:
   - representation of feasible set with ellipsoids: 0.5
   - use of subgradient inequality: 0.5
   - in each step we first intersect ellipsoid with hyper-half-space and next determine smallest ellipsoid that contains the intersection: 1
   - update formulas for center and A matrix of the ellipsoids use the subgradient: 0.5
   - in each step we discard points that are not optimal or that are not feasible (objective/constraint iteration): 1
   - the volume of the ellipsoid gets smaller and smaller as the iteration progresses: 0.5
   - stopping criterion $\|x_k - x^*\| \leq \varepsilon_x$, $\|f(x_k) - f(x^*)\| \leq \varepsilon_f$, $g(x_k) \leq \varepsilon_g$: 1
   - picture: 2

   In case of wrong statements, a penalty of -0.75 applies for each wrong statement.

2 (b) For final solution the ellipsoid algorithm, we also know how far we are maximally away from the true optimum in the variable space: 1
   The size of the LP problems in the cutting-plane and thus also the time needed per iteration increase with each iteration; for the ellipsoid algorithm the computational complexity does not increase as the algorithm progresses: 1

- **Question 2.2**

2 (a) if the problem is characterized as nonconvex, the score for the entire subquestion (a) will be 0
   First of all we transform the maximization problem into a minimization problem:

$$\begin{aligned}\min_{(x,y)\in\mathbb{R}^2} \quad & 10+4x^2+y^2-2xy-50x+20y \\ \text{s.t.} \quad & |x|+|2y| \leq 8 \\ & 1 \leq 2^{(y-1)^2} \leq 2^{16}\end{aligned}$$

   The second constraint can be *equivalently* rewritten as $0 \leq (y-1)^2 \leq 16$ since $2^{(\cdot)}$ is an increasing function. Moreover, the part $0 \leq (y-1)^2$ always hold and can thus be dropped. The other part of the constraint can be *equivalently* rewritten as $|y-1| \leq 4$ or equivalently $-3 \leq y \leq 5$. This then results in the following optimization problem:

$$\begin{aligned}\min_{(x,y)\in\mathbb{R}^2} \quad & 10+4x^2+y^2-2xy-50x+20y \\ \text{s.t.} \quad & |x|+|2y| \leq 8 \\ & -3 \leq y \leq 5\end{aligned}$$

   The objective function can be written as a sum of squares and of linear terms: $3x^2+(x-y)^2-50x+20y+10$ and therefore it is convex: 0.5
   The first constraint involves a sum of norm functions (which are convex), and therefore the constraint is convex: 0.5
   The second constraint involves the norm of a linear function, and therefore this constraint is also convex: 0.75 [this also includes correct simplification of the constraint]
   hence, the given problem is convex: 0.25

12 (b) Note that the constraints can be rewritten as $\pm x \pm 2y \leq 8$, $-3 \leq y \leq 5$, which are linear constraints. So we can indeed apply the gradient projection method.
First we compute the gradient of the objective function: 2

$$\nabla f(x,y) = \begin{bmatrix} 8x - 2y - 50 \\ 2y - 2x + 20 \end{bmatrix}$$

If a sign error is made here or any other computation error later on, marks can still be scored for: use of negative gradient [1], line search equation [1], and the determination of s^* [1], unless s^* is negative or leads to an infeasible point.

In the gradient projection approach the search direction is the **negative** gradient: 1

So we evaluate the negative gradient in the point $(5,0)$: $-\nabla f(5,0) = \begin{bmatrix} 10 & -10 \end{bmatrix}^{\mathrm{T}}$. Since $(5,0)$ is in the interior of the feasible set, no projection is required. So we perform a line search in the direction (10,-10) or equivalently (1,-1), which yield the search line $(x,y) = (5,0) + s(1,-1) = (5+s,-s)$ where s is the step size: 1

If we fill out the values for x and y in the objective function and simplify, we obtain $\bar{f}(s) = 7s^2 - 20s - 140$. We first take the derivative w.r.t. s and put it equal to 0: 1

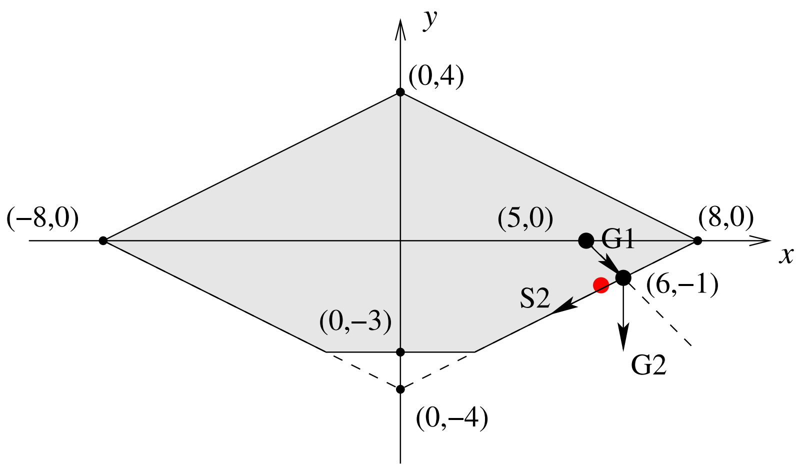
The resulting value is of s is $s^* = \frac{20}{14} = \frac{10}{7}$; as this value is positive it corresponds to the line minimum; however, this step size would create an infeasible point. So we also take the constraints into account: $\pm x \pm 2y \leq 8$, $-3 \leq y \leq 5$, which imply that $-\frac{13}{3} \leq s \leq 1$. So the largest allowed step size is $s = 1$. This yields the point $(x_1, y_1) = (6,-1)$: 3

The negative gradient in $(x_1, y_1) = (6,-1)$ is $-\nabla f(6,-1) = \begin{bmatrix} 0 & -6 \end{bmatrix}^{\mathrm{T}}$, which is pointing away from the feasible set. So we have to project $(0,-6)$ on the active boundary, i.e., on the line $x - 2y = 8$: 1

The directional vector of this line is $(2,1)$. Since $(0,-6)$ points in the direction of decreasing y values, the projected search direction thus is $(-2,-1)$: 1

The search line is thus defined by $(x,y) = (6,-1) + t(-2,-1) = (6-2t, -1-t)$ where t is the step size. If we fill out the values for x and y in the objective function and simplify, we obtain $\bar{f}(t) = 13t^2 - 6t - 153$. We first take the derivative w.r.t. t and put it equal to 0. The resulting value is of t is $t^* = \frac{6}{26} = \frac{3}{13}$; as this value is positive, it corresponds to the line minimum. Moreover, the resulting point $(x_2, y_2) = (\frac{72}{13}, -\frac{16}{13}) = (5 + \frac{7}{13}, -1 - \frac{3}{13}) \approx (5.5, -1.25)$ is feasible: 2

A graphical representation of the procedure above is given by the following picture, where the gray area represents the feasible set and the red dot corresponds to the point (x_2, y_2):



2 (c) We can compute the negative gradient in (x_2, y_2): $\nabla f(x_2, y_2) = \begin{bmatrix} \frac{-42}{13} & \frac{84}{13} \end{bmatrix}^{\mathrm{T}}$ and see that it is orthogonal to the active boundary (which has direction vector $s_2 = \begin{bmatrix} -2 & -1 \end{bmatrix}^{\mathrm{T}}$, as indicated above; it is easy to verify that $(\nabla f(x_2, y_2))^{\mathrm{T}} s_2 = 0$). Moreover, $\nabla f(x_2, y_2)$ points away from the feasible region: 1
This also implies that the projection of the gradient in (x_2, y_2) onto the active boundary will be the zero vector. This in turn implies that the point (x_2, y_2) is a local optimum, and since the problem is convex, also the global optimum: 1

- **Question 2.3**

3 Note that $\mu^{\mathrm{T}} g(x) = 0$ is equivalent to $\sum_{i=1}^{m} \mu_i g_i(x) = 0$ where m is the number of inequality constraints. As $\mu_i \geq 0$ and $g_i(x) \leq 0$, this sum is a sum of negative or zero numbers, which means that it can only be equal to 0 if all the terms $\mu_i g_i(x)$ are equal to 0: 1.5
So if $\mu_i > 0$, we should have $g_i(x) = 0$, i.e., the ith inequality constraint is active (or, in other words, it holds with equality: 1.5