Prof. Dr. Karsten Urban
M. Sc. Lewin Ernst

Winter term 2024/2025
Sheet 11, 28 Points

Numerical Optimization
Solution to exercise sheet

review on 15.01.2024 during the exercise class

1. *(Active set method)*
The active set method is an optimization mehtod for quadratic problems (QP) with (affine) linear equality and inequality constraints of the form

$$\begin{cases} \min\limits_{x} f(x) := \frac{1}{2}x^T A x + a^T x \\ \text{s.t. } Bx = b,\ Cx \leq c, \end{cases} \tag{1}$$

where A is symmetric positive definite on $\ker B$ and B has full rank. The idea is to reduce the problem to a sequence of QPs with only equality constraints. At an iterate $x^{(k)} \in \mathcal{F}$ not being the solution x^*, we seek a *feasible descent direction* $d^{(k)}$ by solving (see (3.6.19))

$$\begin{cases} \min\limits_{d^{(k)}} f(x^{(k)} + d^{(k)}) \\ \text{s.t. } Bd^{(k)} = 0,\ C_j d^{(k)} = 0,\ j \in \mathcal{A}(x^{(k)}) \end{cases}$$

which is equivalent to

$$\begin{cases} \min\limits_{d^{(k)}} \frac{1}{2}(d^{(k)})^T A d^{(k)} + (d^{(k)})^T (Ax^{(k)} + a) \\ \text{s.t. } D^{(k)} d^{(k)} = 0. \end{cases} \tag{2}$$

We know from the lecture, that there can now arise three possibilities.

Case 1: $d^{(k)} = 0$, $\mu^{(k)} \geq 0$

Case 2: $d^{(k)} = 0$, $\mu_j^{(k)} < 0$ for at least on $j \in \mathcal{A}(x^{(k)})$

Case 3: $d^{(k)} \neq 0$ is a feasible direction

a) Derive the KKT system for (1) and (2).

b) Let (1) be convex and let $x^{(k)} \in \mathcal{F}$ be the current iterate. Show that, if for the solution $d^{(k)}$ and the Lagrange multiplier $\mu^{(k)}$ of (2) holds Case 1, then $x^{(k)}$ is a solution of (1).

c) Assume that we are in Case 2 and the Inactivation step has been performed, i.e. we have $d^{(k)} = 0$ and $\mu_j^{(k)} < 0$ for at least one $j \in \mathcal{A}^{(k)}$. Further, we have $\tilde{d}^{(k)}$, $\tilde{\lambda}^{(k)}$ and $\tilde{\mu}^{(k)}$ as the solution of problem (2) w.r.t. the set $\tilde{\mathcal{A}}^{(k)} := \mathcal{A}^{(k)} \setminus \{j\}$ as described in the lecture notes. Show that $\tilde{d}^{(k)}$ is a feasible direction at $x^{(k)}$ by showing $\tilde{d}^{(k)} \in \mathcal{L}(\mathcal{F}, x^{(k)})$.
Hint: Exploit the optimality conditions for $d^{(k)}$ and $\tilde{d}^{(k)}$.

d) Assume everything from c) holds and further assume that $D^{(k)}$ has full rank for all $k \in \mathbb{N}$. Show that $\tilde{d}^{(k)} \neq 0$.

(2 + 4 + 6 + 4 = 16 Points)

Solution:

a) The KKT system for (1) is given by:

$$\begin{aligned} x^T A + a^T + \lambda^T B + \mu^T C &= 0 \\ Bx - b &= 0 \\ \mu &\geq 0 \\ Cx - c &\leq 0 \\ \mu^T(Cx - c) &= 0 \end{aligned}$$

and for (2) we obtain

$$\begin{aligned} (d^{(k)})^T A + (x^{(k)})^T A + a^T + (\lambda^{(k)})^T B + (\mu^{(k)})^T C(x^{(k)}) &= 0 \\ Bd^{(k)} &= 0 \\ C(x^{(k)})d^{(k)} &= 0 \end{aligned}$$

b) We assume now that Case 1 holds, i.e. $d^{(k)} = 0$ and $\mu^{(k)} \geq 0$. We know from the KKT-Theorem for convex problems that if $x^{(k)}$ fulfills the KKT conditions it follows that $x^{(k)}$ is a global solution of (1). Since $d^{(k)} = 0$ is a solution of (2), the KKT system holds for some Lagrange multiplier $\lambda^{(k)}$ and $\mu^{(k)}$. From that and by setting

$$\tilde{\mu}_j^{(k)} := \begin{cases} \mu_j^{(k)}, & j \in \mathcal{A}(x^{(k)}), \\ 0, & \text{else} \end{cases}$$

we deduce

$$(x^{(k)})^T A + a^T + (\lambda^{(k)})^T B + (\tilde{\mu}^{(k)})^T C = 0$$

which is the first equation of the KKT system for (1). Due to $x^{(k)} \in \mathcal{F}$ we have also $Bx^{(k)} - b = 0$ and $Cx^{(k)} - c \leq 0$. The definition of $\tilde{\mu}^{(k)}$ and the assumption of Case 1 gives us $\tilde{\mu}^{(k)} \geq 0$. The last equation $(\tilde{\mu}^{(k)})^T(Cx^{(k)} - c) = 0$ follows also from $\tilde{\mu}_j^{(k)} = 0$ for $j \notin \mathcal{A}(x^{(k)})$ and $C_j x^{(k)} - c_j = 0$ for $j \in \mathcal{A}(x^{(k)})$.

c) We show that $\tilde{d}^{(k)} \in \mathcal{L}(\mathcal{F}, x^{(k)})$, i.e. we have to show $B\tilde{d}^{(k)} = 0$ and $C_j\tilde{d}^{(k)} \leq 0$ for $j \in \mathcal{A}(x^{(k)})$. Due to the constraints of problem (2) we have $B\tilde{d}^{(k)} = 0$ and $C_j\tilde{d}^{(k)} = 0$ for $j \in \tilde{\mathcal{A}}^{(k)}$. The only thing left to prove is $C_{j^*}\tilde{d}^{(k)} \leq 0$ for $j^* := \mathcal{A}^{(k)} \setminus \tilde{\mathcal{A}}^{(k)}$. From the optimality conditions for $d^{(k)}$ and $\tilde{d}^{(k)}$ we deduce

$$\begin{aligned} (x^{(k)})^T A + a^T + (\lambda^{(k)})^T B + (\mu^{(k)})^T C^{(k)} &= 0 \quad \text{and} \\ (\tilde{d}^{(k)})^T A + (x^{(k)})^T A + a^T + (\tilde{\lambda}^{(k)})^T B + (\tilde{\mu}^{(k)})^T \tilde{C}^{(k)} &= 0. \end{aligned}$$

We multiply the equation with $\tilde{d}^{(k)}$ and subtract the resulting equations, from which we get due to the constraints of the QP (2)

$$0 < (\tilde{d}^{(k)})^T A \tilde{d}^{(k)} = \mu_{j^*}^{(k)} C_{j^*}^{(k)} \tilde{d}^{(k)}.$$

Therefore, with the assumption $\mu_{j^*}^{(k)} < 0$ we must have $C_{j^*}^{(k)}\tilde{d}^{(k)} \leq 0$.

d) The derivation from c) holds only if $\tilde{d}^{(k)} \neq 0$. So we assume $\tilde{d}^{(k)} = 0$ and again from the optimality conditions we get

$$\begin{aligned} (x^{(k)})^T A + a^T + (\lambda^{(k)})^T B + (\mu^{(k)})^T C^{(k)} &= 0 \quad \text{and} \\ (x^{(k)})^T A + a^T + (\tilde{\lambda}^{(k)})^T B + (\tilde{\mu}^{(k)})^T \tilde{C}^{(k)} &= 0, \end{aligned}$$

which reduces to

$$(\lambda^{(k)})^T B + (\mu^{(k)})^T C^{(k)} = (\tilde{\lambda}^{(k)})^T B + (\tilde{\mu}^{(k)})^T \tilde{C}^{(k)}.$$

This can be rewritten as

$$(C_{j^*}^{(k)})^T = \frac{1}{\mu_{j^*}^{(k)}} \left[B^T \left(\tilde{\lambda}^{(k)} - \lambda^{(k)} \right) + (\tilde{C}^{(k)})^T \left(\tilde{\mu}^{(k)} - \mu_{\tilde{\mathcal{A}}^{(k)}}^{(k)} \right) \right],$$

i.e. $C_{j^*}^{(k)}$ is a linear combination of the rows of $\tilde{D}^{(k)}$ which is a contradiction to the full rank of $D^{(k)}$.

2. *(Active set method,* MATLAB*)*
In this exercise we want to visualize the iterations of the active set method.

a) Download the MATLAB function `activeset.m` from Moodle and adjust the method such that the `nullspace_method.m` is used to solve the subproblems.

b) Apply the active set method to the problem

$$\begin{aligned} \min_{x \in \mathbb{R}^2} f(x) &= (x_1 - 1)^2 + (x_2 - 2.5)^2 \\ \text{s.t. } x_1 - 2x_2 + 2 &\geq 0 \\ -x_1 - 2x_2 + 6 &\geq 0 \\ -x_1 + 2x_2 + 2 &\geq 0 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

with the initial value $x_0 = (2, 0)^T$ and the associated active set $\mathcal{A}(x_0)$.

c) Plot the contours of f with the constraints and the path of the iterations of the active set method.

(2 + 4 + 6 = 12 Points)

Solution:

a) The adjusted function could look like

```
function [x_opt,lambda_opt,mu_opt,fval,exitflag,its] = ...
                                            activeset_method(A,a,...
                                                             B,b,...
                                                             C,c,...
                                                             x0,tol,maxIter)
% acitve set method for quadratic problems
% Minimize: 0.5*x'*A*x + b'*x
% Subject to: B * x = g, C * x <= r

if nargin<9
    maxIter = 1000;
end

%-- No Inequality Constraints
if size(C,1) == 0
    [x_opt,lambda_opt] = nullspace_method(A,a,B,b);
    mu_opt             = zeros(0,1);
    return
end
```

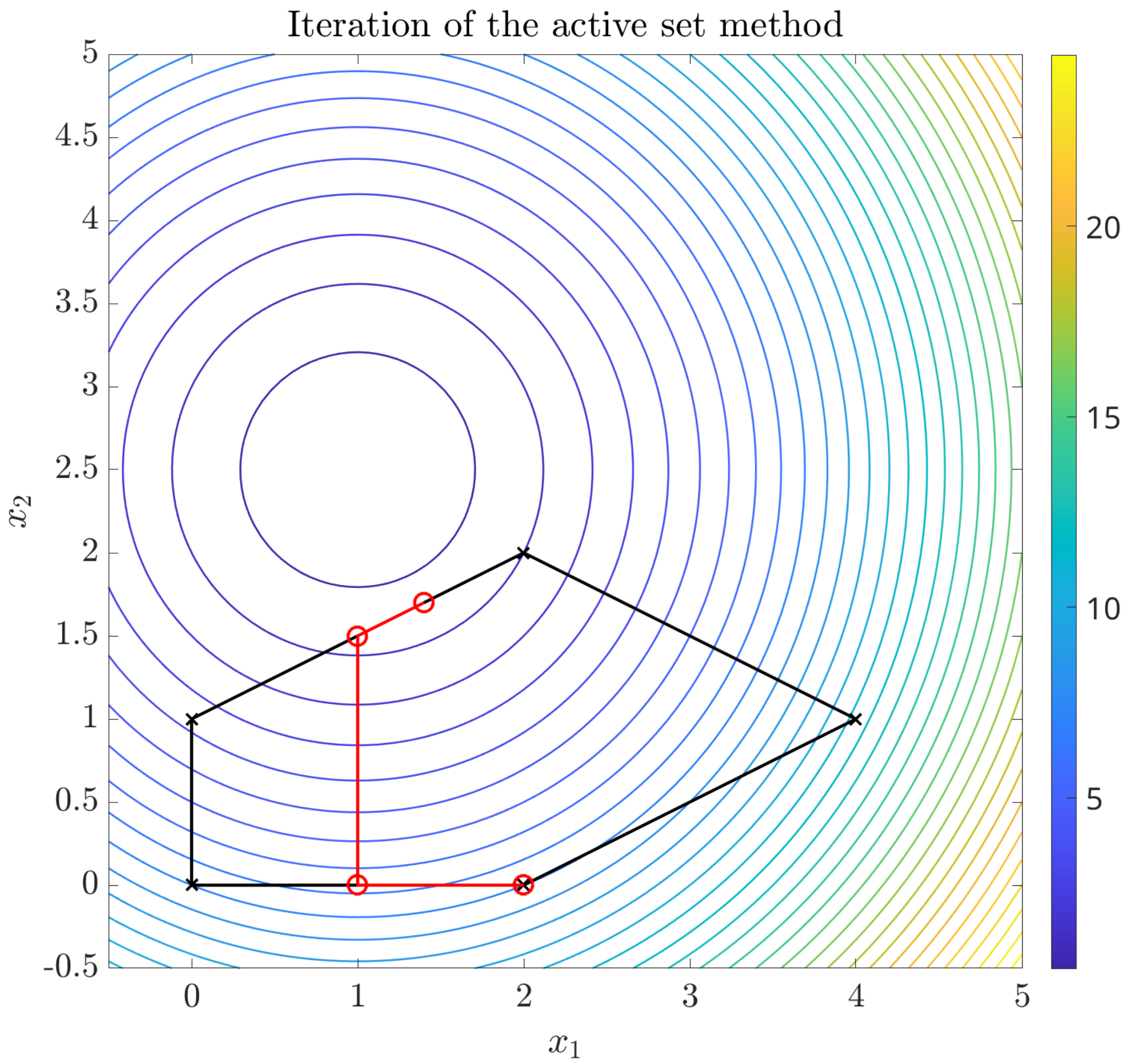
```
%-- Initialization
x(:,1)       = x0;
m            = size(B,1);
lambda       = zeros(m, 1);
mu           = zeros(size(C, 1), 1);
activeSet    = [3,5];
inactiveSet = [1,2,4];
% inactiveSet = 1:size(C, 1);

for iter = 1:maxIter
    %-- nullspace method for feasible descent direction
    ghelp = A*x(:,end) + a;
    Bhelp = [B; C(activeSet, :)];

    if isempty(Bhelp)
        d = -A \ ghelp;
    else
        rhs             = zeros(size(Bhelp, 1),1);
        [d,lambda_mu] = nullspace_method(A,ghelp,Bhelp,rhs);
        lambda          = lambda_mu(1:m);
        if isempty(activeSet)
            mu = [];
        else
            mu = lambda_mu(m+1:end);
        end
    end

    %---------------------------------------
    if norm(d) < tol
        if all(mu >= 0)  %--case 1
            exitflag = 1;
            break;
        else             %-- case 2: inactivation
            [~, idx] = min(mu);
            activeSet(idx) = [];
            continue;
        end
    end

                         %-- case 3
    sigma = 1;           %-- step size determination
    for i = inactiveSet
        if C(i, :) * d > 1e-10
            sigma = min(sigma, (c(i) - C(i, :) * x(:,end)) / (C(i, :) * d));
        end
    end

    x(:,end+1)  = x(:,end) + sigma * d;
    idx          = find(abs(C * x(:,end) - c) < tol)';
    activeSet    = unique([activeSet, idx]);
    inactiveSet = setdiff(1:size(C, 1), activeSet);
end
mu_opt              = zeros(size(C, 1),1);
mu_opt(activeSet) = mu;
lambda_opt          = lambda;
x_opt = x;
fval  = 0.5 * x(:,end)' * A * x(:,end) + a' * x(:,end);
its   = iter;
if iter == maxIter
    exitflag = 0;
end
```

```
end
```

b) The script could look like

```
clear, close all
clc

% define the problem by defining the matrices
A = 2*eye(2,2);
a = -[2; 5];
B = [];
b = [];
C = -[1, -2; -1, -2; -1, 2; 1, 0; 0, 1];
c = [2; 6; 2; 0; 0];
f_opt = @(x) x'*A*x + a'*x;

% solver options
tol     = 1e-10;
maxIter = 50;
x0      = [2;0];

[x_opt,lambda,mu,fval,exitflag,its] = activeset_method(A,a,B,b,C,c,x0,tol,maxIter);

% print the function and the solution path
x = linspace(-0.5,5,96);
y = linspace(-0.5,5,96);
[xx,yy] = meshgrid(x,y);
ff = (xx-1).^2 + (yy-2.5).^2;
levels = -1:0.75:50;
figure(1)
contour(x,y,ff,levels,LineWidth=1.2),
colorbar
axis([-0.5 5 -0.5 5])
axis square
hold on
% plot constraints
plot([0, 2, 4, 2, 0, 0],[1, 2, 1, 0, 0, 1],'kx-', MarkerSize=10, LineWidth=2);
% plot iterations
plot(x_opt(1,:), x_opt(2,:), 'r-o', MarkerSize=12, LineWidth=2)
xlabel("$x_1$", Interpreter="latex")
ylabel("$x_2$", Interpreter="latex")
title("Iteration of the active set method", Interpreter="latex")
set(gca, "FontSize", 22, "TickLabelInterpreter", "latex")
hold off
```

c) The Plot could look like


Iteration of the active set method
x_1
x_2
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
0
1
2
3
4
5
20
15
10
5