Stochastic optimization algorithms
Lecture 7, 20200915

Classical optimization methods and evolutionary algorithms: Problem-solving

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

CHALMERS

Information (1)

- The correction of the Introductory programming problem (IPP) is ongoing.
- We will upload the results no later than Thursday (17th).
- I will list frequent comments and discuss the IPP on Friday.
- Comments are given for *your* benefit, to minimize the risk of unnecessary point loss in the correction of HP1.
- Make sure to read the IPP comments before submitting HP1 (making any necessary code adjustments).
- Unless clearly stated, you do not need to resubmit the IPP.

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Information (2)

- Brief solutions to some problems in Chapter 2 have been uploaded to the course web page.
- See the module *Miscellaneous*.

Today's learning goals

- After this lecture you should be able to
  - Solve problems 2.12 and 2.13 in the book
  - Solve problems 3.2 and 3.9 in the book

Problem 2.12

- Find minimum of ${x_1}^2 + 2{x_2}^2$ subject to the equality constraint $2{x_1}^2 + x_2 - 3 = 0$.
- Use the penalty method.

Problem 2.12

- With this function

$$f(x_1, x_2) = x_1^2 + 2x_2^2 \quad (1)$$

- and the equality constraint

$$2x_1^3 + x_2 = 3 \quad (2)$$

- one can write the function $f_p(\boldsymbol{x};\mu)$ as

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu\left(2x_1^2 + x_2 - 3\right)^2 \quad (3)$$

Problem 2.12

- With this function

$$f(x_1, x_2) = x_1^2 + 2x_2^2 \tag{1}$$

- and the equality constraint

$$2x_1^3 + x_2 = 3 \tag{2}$$

- one can write the function $f_p(\boldsymbol{x}; \mu)$ as

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu \underbrace{\left(2x_1^2 + x_2 - 3\right)}_{h(x_1, x_2)}{}^2 \tag{3}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12

- With this function

$$f(x_1, x_2) = x_1^2 + 2x_2^2 \tag{1}$$

- and the equality constraint

$$2x_1^3 + x_2 = 3 \tag{2}$$

- one can write the function $f_p(\boldsymbol{x}; \mu)$ as

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu \underbrace{\left(2x_1^2 + x_2 - 3\right)}_{h(x_1, x_2)}{}^2 \tag{3}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu\left(2x_1^2 + x_2 - 3\right)^2 \tag{3}$$

- Computing the first gradient component and setting it to zero, we get

$$\frac{\partial f}{\partial x_1} = 2x_1 + 8\mu x_1\left(2x_1^2 + x_2 - 3\right) = 0 \tag{4}$$

- ..where, for the first equation, we have made use of

$$\frac{\partial(2x_1^2)}{\partial x_1} = 4x_1$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu\left(2x_1^2 + x_2 - 3\right)^2 \tag{3}$$

- Computing the first gradient component and setting it to zero, we get

$$\frac{\partial f}{\partial x_1} = 2x_1 + 8\mu x_1\left(2x_1^2 + x_2 - 3\right) = 0 \tag{4}$$

- ..where, for the first equation, we have made use of

$$\frac{\partial(2x_1^2)}{\partial x_1} = 4x_1$$

Problem 2.12

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu\left(2x_1^2 + x_2 - 3\right)^2 \quad (3)$$

- Computing the first gradient component and setting it to zero, we get

$$\frac{\partial f}{\partial x_1} = 2x_1 + 8\mu x_1\left(2x_1^2 + x_2 - 3\right) = 0 \quad (4)$$

- ..where, for the first equation, we have made use of

$$\frac{\partial(2x_1^2)}{\partial x_1} = 4x_1$$

Problem 2.12

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu\left(2x_1^2 + x_2 - 3\right)^2 \qquad (3)$$

- Computing the second gradient component and setting it to zero, we get

$$\frac{\partial f}{\partial x_2} = 4x_2 + 2\mu\left(2x_1^2 + x_2 - 3\right) = 0 \qquad (5)$$

Problem 2.12

$$f_p(x_1, x_2; \mu) = x_1^2 + 2x_2^2 + \mu\left(2x_1^2 + x_2 - 3\right)^2 \quad (3)$$

- Computing the second gradient component and setting it to zero, we get

$$\frac{\partial f}{\partial x_2} = 4x_2 + 2\mu\left(2x_1^2 + x_2 - 3\right) = 0 \quad (5)$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12

- Then, to eliminate the second term in Eqs. (4) and (5), take Eq. (5) multiplied by $4x_1$ and subtract Eq. (4):

$$\frac{\partial f}{\partial x_2} = 4x_2 + 2\mu\left(2x_1^2 + x_2 - 3\right) = 0 \quad (5)$$

$$\frac{\partial f}{\partial x_1} = 2x_1 + 8\mu x_1\left(2x_1^2 + x_2 - 3\right) = 0 \quad (4)$$

- to obtain

$$16x_1x_2 - 2x_1 = 0 \quad (6)$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12

- Then, to eliminate the second term in Eqs. (4) and (5), take Eq. (5) multiplied by $4x_1$ and subtract Eq. (4):

$4x_1$

$$\frac{\partial f}{\partial x_2} = 4x_2 + 2\mu\left(2x_1^2 + x_2 - 3\right) = 0 \qquad (5)$$

-1

$$\frac{\partial f}{\partial x_1} = 2x_1 + 8\mu x_1\left(2x_1^2 + x_2 - 3\right) = 0 \qquad (4)$$

- to obtain

$$16x_1x_2 - 2x_1 = 0 \qquad (6)$$

Problem 2.12

- ...from which it is obvious that

$$x_2 = \frac{1}{8} \tag{7}$$

- ...unless $x_1 = 0$, in which case one gets:

$$x_1 = 0 \Rightarrow x_2 = 3 \Rightarrow f(x_1, x_2) = 18 \tag{8}$$

- we'll keep this possibility in mind, but for now we will proceed with the first case ($x_1 \neq 0$).

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12

- Then we get (from Eq. (5), inserting $x_2 = 1/8$)

$$\frac{4}{8} + 2\mu\left(2x_1^2 + \frac{1}{8} - 3\right) = 0 \tag{9}$$

- so that ...

$$\frac{1}{2} + 2\mu\left(2x_1^2 - \frac{23}{8}\right) = 0 \tag{10}$$

- ..and then

$$x_1^2 = -\frac{1}{8\mu} + \frac{23}{16} \tag{11}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12

- ..from which

$$x_1 = \pm\sqrt{-\frac{1}{8\mu} + \frac{23}{16}} \qquad (12)$$

- Thus, letting μ tend to infinity, the result becomes

$$x_1 = \pm\frac{\sqrt{23}}{4} \qquad (13)$$

Problem 2.12

- Therefore:

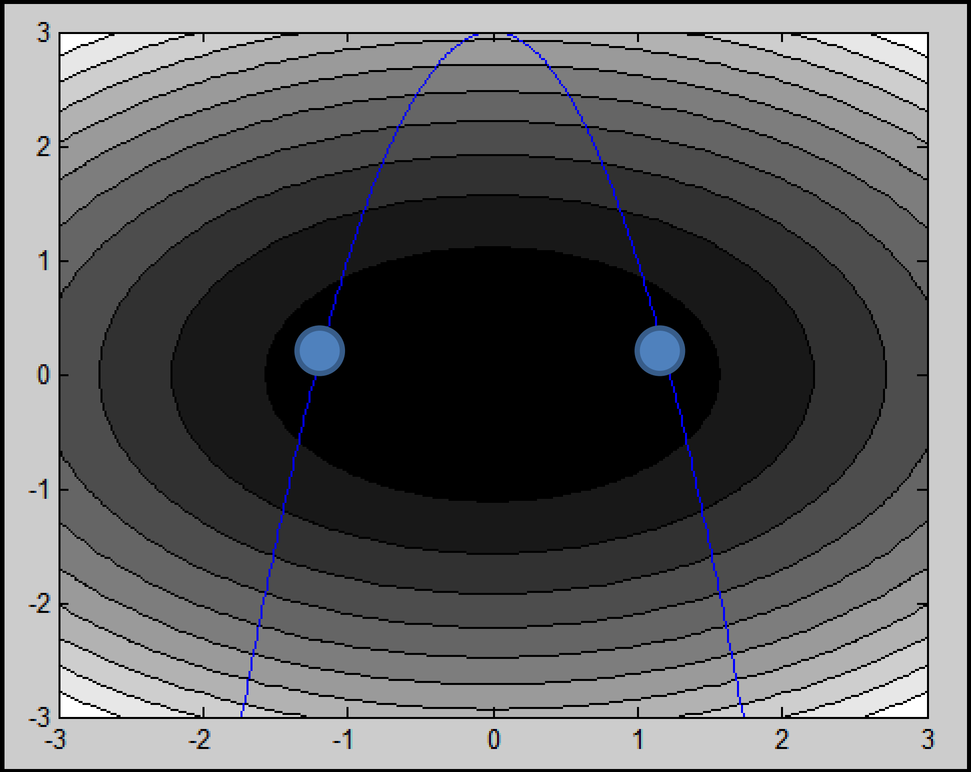
$$(x_1, x_2)^{\mathrm{T}} = \left(\pm\frac{\sqrt{23}}{4}, \frac{1}{8}\right)^{\mathrm{T}} \tag{14}$$

- The function value at these points is:

$$f(x_1, x_2) = \frac{23}{16} + \frac{2}{64} = \frac{47}{32}. \tag{15}$$

- So, the minimum occurs at the points given in Eq. (14), where the function takes the value 47/32.

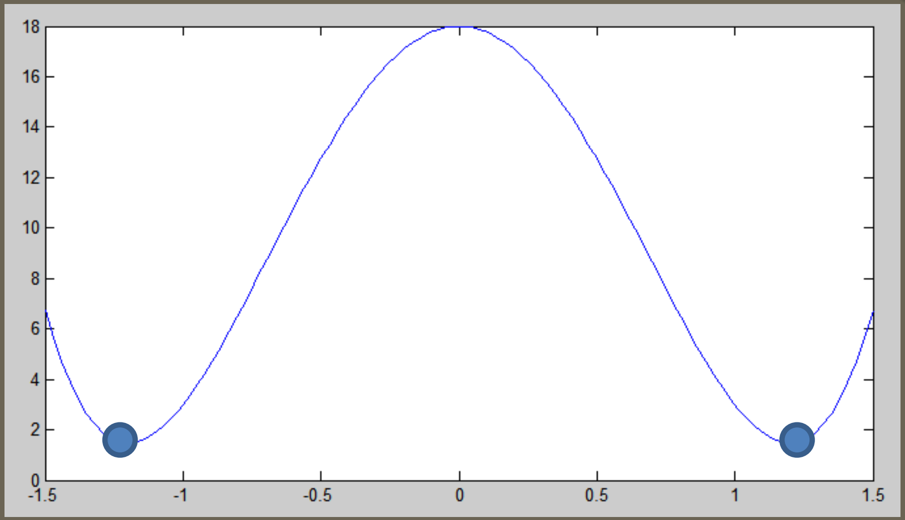
Problem 2.12



Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.12, alternative approach

- Eliminate $x_2 \Rightarrow \mathrm{g}(x_1) = {x_1}^2 + 2(3 - 2{x_1}^2)^2$.
- Solve without penalty method (e.g. Newton-Raphson or analytically ($\mathrm{g}'(x_1) = 0 \Rightarrow \ldots$)



Problem 2.13

- Find the *maximum* value of $\mathrm{f}(x_1, x_2) = 2{x_1}^2 - 4x_1 + {x_2}^2 + 2x_2$ subject to the inequality constraint $2{x_1}^2 + {x_2}^2 \leq 12$.
- Analytical method:
  - Find stationary points in the interior of the set S defined by the constraint.
  - Find stationary points on the boundary ∂S.
  - Check the points one by one.

Problem 2.13

- With the function

$$f(x_1, x_2) = 2x_1^2 - 4x_1 + x_2^2 + 2x_2 \quad (16)$$

- over the set S defined by

$$2x_1^2 + x_2^2 \leq 12. \quad (17)$$

- we get for the interior of S:

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) = 0 \Rightarrow \quad (18)$$

- ...so that

$$4x_1 - 4 = 0 \quad (19)$$

$$2x_2 + 2 = 0 \quad (20)$$

Problem 2.13

- ...which gives us one point:

$$P_1 = (1, -1)^{\mathrm{T}}. \qquad (21)$$

Problem 2.13

- For the boundary ∂S we can use the Lagrange multiplier method.
- On the boundary, we have

$$2x_1^2 + x_2^2 - 12 = 0$$

- ..and therefore

$$L(x_1, x_2, \lambda) = 2x_1^2 - 4x_1 + x_2^2 + 2x_2 + \lambda\left(2x_1^2 + x_2^2 - 12\right) \qquad (22)$$

Problem 2.13

- so that ...

$$\frac{\partial L}{\partial x_1} = 4x_1 - 4 + 4\lambda x_1 = 0 \tag{23}$$

$$\frac{\partial L}{\partial x_2} = 2x_2 + 2 + 2\lambda x_2 = 0 \tag{24}$$

$$\frac{\partial L}{\partial \lambda} = 2x_1^2 + x_2^2 - 12 = 0 \tag{25}$$

- We then get

$$x_1(\lambda + 1) = 1 \Rightarrow x_1 - 1 = \lambda x_1 \tag{26}$$

$$x_2(\lambda + 1) = -1 \Rightarrow x_2 + 1 = -\lambda x_2 \tag{27}$$

Problem 2.13

- Here we will assume that x_1 and x_2 are both different from 0. If not, then we get:

$$x_1 = 0 \Rightarrow x_2 = \pm\sqrt{12} \Rightarrow f(x_1, x_2) = 12 \pm 2\sqrt{12} < 24 \qquad (28)$$

$$x_2 = 0 \Rightarrow x_1 = \pm\sqrt{6} \Rightarrow f(x_1, x_2) = 12 \pm 4\sqrt{6} < 24 \qquad (29)$$

- We will keep this possibilities in mind for later ...
- Proceeding now from Eqs. (26)-(27), assuming that x_1 and x_2 are both different from 0:

$$-\lambda = \frac{x_1 - 1}{x_1} = \frac{x_2 + 1}{x_2} \Rightarrow x_2 = -x_1 \qquad (30)$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.13

- Inserting this into the constraint equation, we get

$$2x_1^2 + (-x_1)^2 = 3x_1^2 = 12 \Rightarrow x_1 = \pm 2. \tag{31}$$

- since $x_2 = -x_1$ we get two solutions:

$$P_2 = (2, -2)^{\mathrm{T}},\ P_3 = (-2, 2)^{\mathrm{T}} \tag{32}$$

- Computing the value of f at the three points, we get

$$f(P_1) = -3,\ f(P_2) = 0\ f(P_3) = 24. \tag{33}$$

Problem 2.13

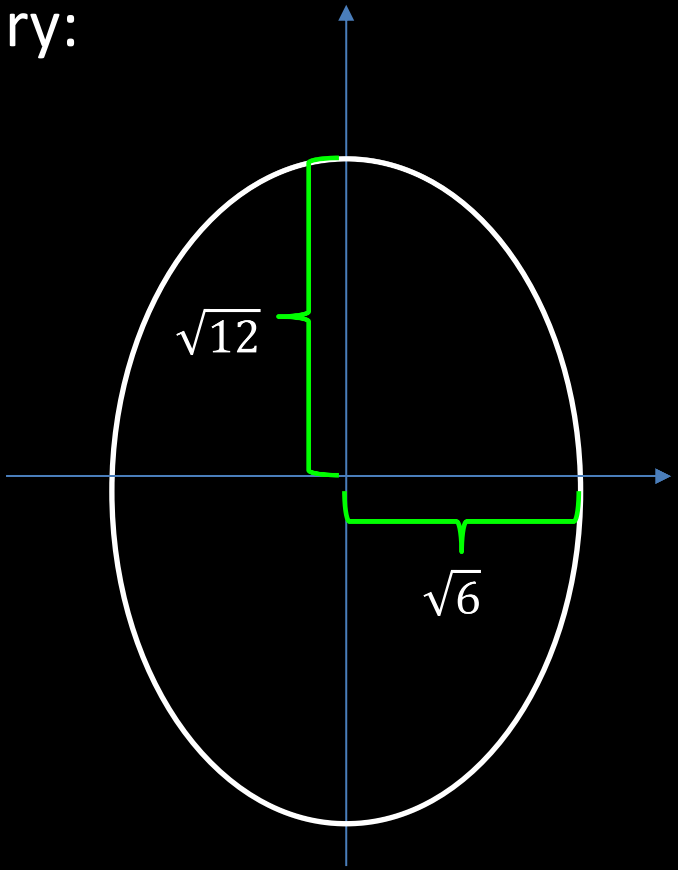
- Thus the maximum is at P_3 where the function takes the value 24.
- Note that if the problem had been to find the *minimum*, we could have made use of the fact that f is convex, and so is the inequality constraint, meaning that the point P_1 found in the interior could directly be identified as the minimum... (no need to check the boundary, in that case).

Problem 2.13

- Alternative approach for the boundary:

$$2x_1^2 + x_2^2 = 12 \Rightarrow \left(\frac{x_1}{\sqrt{6}}\right)^2 + \left(\frac{x_2}{\sqrt{12}}\right)^2 = 1 \qquad (34)$$

$$x_1 = \sqrt{6}\cos t,\ \ x_2 = \sqrt{12}\sin t \qquad (35)$$



Problem 2.13

- The equation for the boundary becomes:

$$
\begin{aligned}
k(t) &= f(x_1(t)), x_2(t)) = 12\cos^2 t - 4\sqrt{6}\cos t + 12\sin^2 t + 2\sqrt{12}\sin t = \\
&= 12 - 4\sqrt{6}\cos t + 2\sqrt{12}\sin t \qquad (36)
\end{aligned}
$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.13

- The equation for the boundary becomes:

$$
\begin{aligned}
k(t) &= f(x_1(t)), x_2(t)) = 12\cos^2 t - 4\sqrt{6}\cos t + 12\sin^2 t + 2\sqrt{12}\sin t = \\
&= 12 - 4\sqrt{6}\cos t + 2\sqrt{12}\sin t \qquad (36)
\end{aligned}
$$

$= 2x_1^2$ etc.

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 2.13

- The equation for the boundary becomes:

$$\begin{aligned} k(t) &= f(x_1(t)), x_2(t)) = 12\cos^2 t - 4\sqrt{6}\cos t + 12\sin^2 t + 2\sqrt{12}\sin t = \\ &= 12 - 4\sqrt{6}\cos t + 2\sqrt{12}\sin t \end{aligned} \tag{36}$$

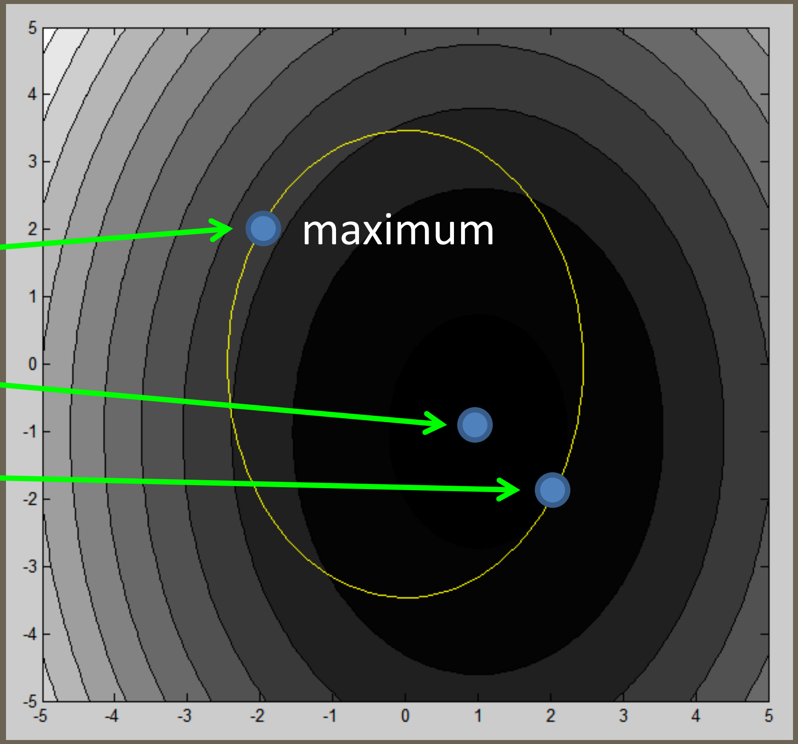
$$k'(t) = 4\sqrt{6}\sin t + 2\sqrt{12}\cos t = 0 \Rightarrow \tan t = -\frac{1}{\sqrt{2}} \tag{37}$$

Problem 2.13

$(x_1, x_2)^T = (-2, 2)^T, \mathrm{f} = 24$

$(x_1, x_2)^T = (1, -1)^T, \mathrm{f} = -3$

$(x_1, x_2)^T = (2, -2)^T, \mathrm{f} = 0$



Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde
CHALMERS

Today’s learning goals

- After this lecture you should be able to
  - Solve problems 2.12 and 2.13 in the book ✓
  - Solve problems 3.2 and 3.9 in the book

Problem 3.2

- Population of five individuals
- Fitness values: $F_1 = 5, F_2 = 7, F_3 = 8, F_4 = 10, F_5 = 15.$
- Find the probability of selecting individual 4 (in a single selection step) with
  - Roulette-wheel selection (RWS)
  - Tournament selection ($P_{\text{tour}} = 0.75$)
  - Roulette-wheel selection with ranking from 1 to 10.

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

CHALMERS

Problem 3.2: Roulette-wheel selection

- $\sum_{i=1}^{5} F_i$ = 5+7+8+10+15 = 45
- Probability of selecting individual 4 = $F_4/\sum_{i=1}^{5} F_i$ = 10/45 = 2/9 ≈ 0.2222.

Problem 3.2: Tournament selection

- 25 possible pairs, of which 9 include individual 4.
- All pairs equally likely.
- Individual 4 is the better individual in 6 of the pairs, and the worse individual in two of the pairs.
- $p_4 = \frac{1}{25}(1 + 6 * 0.75 + 2 * 0.25) = 0.24$

(1,1)			(1,4)	
			(2,4)	
			(3,4)	
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)
			(5,4)	(5,5)

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.2: Tournament selection

- 25 possible pairs, of which 9 include individual 4.
- All pairs equally likely.
- Individual 4 is the better individual in 6 of the pairs, and the worse individual in two of the pairs.
- $p_4 = \frac{1}{25}(1 + 6 * 0.75 + 2 * 0.25) = 0.24$

(1,1)			(1,4)	
			(2,4)	
			(3,4)	
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)
			(5,4)	(5,5)

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde
CHALMERS

Problem 3.2: Tournament selection

- 25 possible pairs, of which 9 include individual 4.
- All pairs equally likely.
- Individual 4 is the better individual in 6 of the pairs, and the worse individual in two of the pairs.
- $p_4 = \frac{1}{25}(1 + 6 * 0.75 + \boxed{2 * 0.25}) = 0.24$

(1,1)			(1,4)	
			(2,4)	
			(3,4)	
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)
			(5,4)	(5,5)

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.2: Tournament selection

- 25 possible pairs, of which 9 include individual 4.
- All pairs equally likely.
- Individual 4 is the better individual in 6 of the pairs, and the worse individual in two of the pairs.
- $p_4 = \frac{1}{25}(1 + 6 * 0.75 + 2 * 0.25) = 0.24$

(1,1)			(1,4)	
			(2,4)	
			(3,4)	
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)
			(5,4)	(5,5)

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.2: RWS with ranking

- The individuals are already sorted in increasing order of fitness, so the first individual is the worst, the second is the second worst etc.
- Ranked fitness values, range 1 to 10:
- $F_1^{rank} = 1$ (worst individual), $F_5^{rank} = 10$ (best individual)
- $F_3^{rank} = (1+10)/2 = 5.5$
- $F_2^{rank} = 3.25$, $F_4^{rank} = 7.75$
- $\sum_{i=1}^{5} F_i^{rank} = 1 + 3.25 + 5.5 + 7.75 + 10 = 27.5$
- Probability of selecting individual 4 = 7.75/27.5 = 0.2818

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.2: RWS with ranking

- The individuals are already sorted in increasing order of fitness, so the first individual is the worst, the second is the second worst etc.
- Ranked fitness values, range 1 to 10:
- $F_1^{rank} = 1$ (worst individual), $F_5^{rank} = 10$ (best individual)
- $F_3^{rank} = (1+10)/2 = 5.5$
- $F_2^{rank} = 3.25$, $F_4^{rank} = 7.75$
- $\sum_{i=1}^{5} F_i^{rank} = 1 + 3.25 + 5.5 + 7.75 + 10 = 27.5$
- Probability of selecting individual 4 = 7.75/27.5 = 0.2818

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.2: RWS with ranking

- The individuals are already sorted in increasing order of fitness, so the first individual is the worst, the second is the second worst etc.
- Ranked fitness values, range 1 to 10:
- $F_1^{rank} = 1$ (worst individual), $F_5^{rank} = 10$ (best individual)
- $F_3^{rank} = (1+10)/2 = 5.5$
- $F_2^{rank} = 3.25$, $F_4^{rank} = 7.75$
- $\sum_{i=1}^{5} F_i^{rank} = 1 + 3.25 + 5.5 + 7.75 + 10 = 27.5$
- Probability of selecting individual 4 = 7.75/27.5 = 0.2818

Problem 3.2: RWS with ranking

- The individuals are already sorted in increasing order of fitness, so the first individual is the worst, the second is the second worst etc.
- Ranked fitness values, range 1 to 10:
- $F_1^{rank} = 1$ (worst individual), $F_5^{rank} = 10$ (best individual)
- $F_3^{rank} = (1+10)/2 = 5.5$
- $F_2^{rank} = 3.25$, $F_4^{rank} = 7.75$
- $\sum_{i=1}^{5} F_i^{rank} = 1 + 3.25 + 5.5 + 7.75 + 10 = 27.5$
- Probability of selecting individual 4 = 7.75/27.5 = 0.2818

Problem 3.2: RWS with ranking

- The individuals are already sorted in increasing order of fitness, so the first individual is the worst, the second is the second worst etc.
- Ranked fitness values, range 1 to 10:
- $F_1^{rank} = 1$ (worst individual), $F_5^{rank} = 10$ (best individual)
- $F_3^{rank} = (1+10)/2 = 5.5$
- $F_2^{rank} = 3.25$, $F_4^{rank} = 7.75$
- $\sum_{i=1}^{5} F_i^{rank} = 1 + 3.25 + 5.5 + 7.75 + 10 = 27.5$
- Probability of selecting individual 4 = 7.75/27.5 = 0.2818

Problem 3.2: RWS with ranking

- The individuals are already sorted in increasing order of fitness, so the first individual is the worst, the second is the second worst etc.
- Ranked fitness values, range 1 to 10:
- $F_1^{rank} = 1$ (worst individual), $F_5^{rank} = 10$ (best individual)
- $F_3^{rank} = (1+10)/2 = 5.5$
- $F_2^{rank} = 3.25$, $F_4^{rank} = 7.75$
- $\sum_{i=1}^{5} F_i^{rank} = 1 + 3.25 + 5.5 + 7.75 + 10 = 27.5$
- Probability of selecting individual 4 = 7.75/27.5 = 0.2818

Problem 3.9

- Infinite-population GA.
- Fitness function: A function of unitation, namely $f(j) = j(m - j)$.
- Find
  - (a) The average fitness in the first generation.
  - (b) The probability distribution in the second generation.
  - (c) The average fitness in the second generation.

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Average fitness:

$$\overline{F}_1 = \sum_{j=0}^{m} F(j)p_1(j) \tag{39}$$

- Initial probability distribution (randomly initialized binary strings):

$$P_1(j) = 2^{-m}\binom{m}{j} \tag{38}$$

Problem 3.9

- Average fitness:

$$\overline{F}_1 = \sum_{j=0}^{m} F(j) p_1(j) \qquad (39)$$

- Initial probability distribution (randomly initialized binary strings):

$$P_1(j) = 2^{-m} \binom{m}{j} \qquad (38)$$

Problem 3.9

- Average fitness:

$$
\begin{aligned}
\overline{F}_1 &= 2^{-m}\sum_{j=0}^{m} j(m-j)\binom{m}{j} = \\
&= 2^{-m} m \sum_{j=0}^{m} j\binom{m}{j} - 2^{-m}\sum_{j=0}^{m} j^2\binom{m}{j} = \\
&= 2^{-m} m \left(m2^{m-1}\right) - 2^{-m}\left(m(m+1)2^{m-2}\right) = \\
&= \frac{m^2}{2} - \frac{m(m+1)}{4} = \frac{m(m-1)}{4} \qquad (40)
\end{aligned}
$$

Problem 3.9

- Average fitness:

$$
\begin{aligned}
\overline{F}_1 &= 2^{-m} \sum_{j=0}^{m} j(m-j) \binom{m}{j} = \\
&= 2^{-m} m \sum_{j=0}^{m} j \binom{m}{j} - 2^{-m} \sum_{j=0}^{m} j^2 \binom{m}{j} = \\
&= 2^{-m} m \left(m 2^{m-1} \right) - 2^{-m} \left(m(m+1) 2^{m-2} \right) = \\
&= \frac{m^2}{2} + - \frac{m(m+1)}{4} = \frac{m(m-1)}{4} \qquad (40)
\end{aligned}
$$

Eq. (B17)

Eq. (B23)

Problem 3.9

- Average fitness:

$$
\begin{aligned}
\overline{F}_1 &= 2^{-m}\sum_{j=0}^{m} j(m-j)\binom{m}{j} = \\
&= 2^{-m} m \sum_{j=0}^{m} j\binom{m}{j} - 2^{-m}\sum_{j=0}^{m} j^2\binom{m}{j} = \\
&= 2^{-m} m \left(m2^{m-1}\right) - 2^{-m}\left(m(m+1)2^{m-2}\right) = \\
&= \frac{m^2}{2} \quad - \frac{m(m+1)}{4} = \frac{m(m-1)}{4} \qquad (40)
\end{aligned}
$$

Problem 3.9

- Probability distribution in the second generation:

$$
\begin{aligned}
p_2(j) &= \frac{F(j)p_1(j)}{\sum_{j=0}^{m} F(j)p_1(j)} \equiv \frac{F(j)p_1(j)}{\overline{F}_1} = \\
&= \frac{j(m-j)2^{-m}\binom{m}{j}}{\frac{m(m-1)}{4}} = 2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \qquad (41)
\end{aligned}
$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Probability distribution in the second generation:

$$
\begin{aligned}
p_2(j) &= \frac{F(j)p_1(j)}{\sum_{j=0}^{m} F(j)p_1(j)} \equiv \frac{F(j)p_1(j)}{\overline{F}_1} = \\
&= \frac{j(m-j)2^{-m}\binom{m}{j}}{\frac{m(m-1)}{4}} = 2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \qquad (41)
\end{aligned}
$$

Problem 3.9

- Probability distribution in the second generation:

$$
\begin{aligned}
p_2(j) &= \frac{F(j)p_1(j)}{\sum_{j=0}^{m} F(j)p_1(j)} \equiv \frac{F(j)p_1(j)}{\overline{F}_1} = \\
&= \frac{j(m-j)2^{-m}\binom{m}{j}}{\frac{m(m-1)}{4}} = 2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \qquad (41)
\end{aligned}
$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Probability distribution in the second generation:

$$p_2(j) = \frac{F(j)p_1(j)}{\sum_{j=0}^{m} F(j)p_1(j)} \equiv \frac{F(j)p_1(j)}{\overline{F}_1} =$$

$$= \frac{j(m-j)2^{-m}\binom{m}{j}}{\frac{m(m-1)}{4}} = 2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \quad (41)$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Probability distribution in the second generation:

$$
\begin{aligned}
p_2(j) &= \frac{F(j)p_1(j)}{\sum_{j=0}^{m} F(j)p_1(j)} \equiv \frac{F(j)p_1(j)}{\overline{F}_1} = \\
&= \frac{j(m-j)2^{-m}\binom{m}{j}}{\frac{m(m-1)}{4}} = 2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \qquad (41)
\end{aligned}
$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Average fitness in the second generation

$$\begin{aligned}\overline{F}_2 &= \sum_{j=0}^{m} F(j)p_2(j) = \sum_{j=0}^{m} j(m-j)2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \\ &= 2^{2-m}\frac{1}{m(m-1)}\sum_{j=0}^{m} j^2(m-j)^2\binom{m}{j} \qquad (42)\end{aligned}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Average fitness in the second generation

$$
\begin{aligned}
\overline{F}_2 &= \sum_{j=0}^{m} F(j)p_2(j) = \sum_{j=0}^{m} j(m-j)2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \\
&= 2^{2-m}\frac{1}{m(m-1)}\sum_{j=0}^{m} j^2(m-j)^2\binom{m}{j} \qquad (42)
\end{aligned}
$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Method 1: Brute force – compute the required sums, namely $\sum_{j=0}^{m} j^3 \binom{m}{j}$ and $\sum_{j=0}^{m} j^4 \binom{m}{j}$.

Problem 3.9

- Average fitness in the second generation

$$\begin{aligned}\overline{F}_2 &= \sum_{j=0}^{m} F(j)p_2(j) = \sum_{j=0}^{m} j(m-j)2^{2-m}\frac{j(m-j)}{m(m-1)}\binom{m}{j} \\ &= 2^{2-m}\frac{1}{m(m-1)}\sum_{j=0}^{m} j^2(m-j)^2\binom{m}{j} \qquad (42)\end{aligned}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Method 1: Brute force – compute the required sums, namely $\sum_{j=0}^{m} j^3 \binom{m}{j}$ and $\sum_{j=0}^{m} j^4 \binom{m}{j}$.

Problem 3.9

- Method 2: Using the binomial theorem directly:

$$(a+b)^m = \sum_{j=0}^{m} a^j b^{m-j} \binom{m}{j} \qquad (43)$$

- We need:

$$\sum_{j=0}^{m} j^2 (m-j)^2 \binom{m}{j}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Take the partial derivative with respect to *a*

$$m(a+b)^{m-1} = \sum_{j=0}^{m} j a^{j-1} b^{m-j} \binom{m}{j} \tag{44}$$

- … then multiply by *a*:

$$ma(a+b)^{m-1} = \sum_{j=0}^{m} j a^{j} b^{m-j} \binom{m}{j} \tag{45}$$

Problem 3.9

- Take the partial derivative with respect to *a*

$$m(a+b)^{m-1} = \sum_{j=0}^{m} j a^{j-1} b^{m-j} \binom{m}{j} \tag{44}$$

- … then multiply by *a*:

$$ma(a+b)^{m-1} = \sum_{j=0}^{m} j a^{j} b^{m-j} \binom{m}{j} \tag{45}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Now take the partial derivative with respect to *b*

$$m(m-1)a(a+b)^{m-2} = \sum_{j=0}^{m} j(m-j)a^j b^{m-j-1} \binom{m}{j} \tag{46}$$

- … then multiply by *b*:

$$m(m-1)ab(a+b)^{m-2} = \sum_{j=0}^{m} j(m-j)a^j b^{m-j} \binom{m}{j} \tag{47}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Now take the partial derivative with respect to *b*

$$m(m-1)a(a+b)^{m-2} = \sum_{j=0}^{m} j(m-j)a^j b^{m-j-1}\binom{m}{j} \tag{46}$$

- … then multiply by *b*:

$$m(m-1)ab(a+b)^{m-2} = \sum_{j=0}^{m} j(m-j)a^j b^{m-j}\binom{m}{j} \tag{47}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Now again take the partial derivative with respect to *a*

$$m(m-1)b(a+b)^{m-2}+m(m-1)(m-2)ab(a+b)^{m-3}=\sum_{j=0}^{m} j^2(m-j)a^{j-1}b^{m-j}\binom{m}{j} \tag{48}$$

- … then multiply by *a* once more:

$$m(m-1)ba(a+b)^{m-2}+m(m-1)(m-2)a^2b(a+b)^{m-3}=\sum_{j=0}^{m} j^2(m-j)a^{j}b^{m-j}\binom{m}{j} \tag{49}$$

Problem 3.9

- Now again take the partial derivative with respect to *a*

$$m(m-1)b(a+b)^{m-2}+m(m-1)(m-2)ab(a+b)^{m-3}=\sum_{j=0}^{m} j^2(m-j)a^{j-1}b^{m-j}\binom{m}{j} \tag{48}$$

- … then multiply by *a* once more:

$$m(m-1)ba(a+b)^{m-2}+m(m-1)(m-2)a^2b(a+b)^{m-3}=\sum_{j=0}^{m} j^2(m-j)a^{j}b^{m-j}\binom{m}{j} \tag{49}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Now again take the partial derivative with respect to *b*

$$m(m-1)a(a+b)^{m-2} + m(m-1)(m-2)ba(a+b)^{m-3} + m(m-1)(m-2)a^2(a+b)^{m-3} + \\ + \quad m(m-1)(m-2)(m-3)a^2b(a+b)^{m-4} = \sum_{j=0}^{m} j^2(m-j)^2 a^j b^{m-j-1} \binom{m}{j} \qquad (50)$$

- … then multiply by *b* once more:

$$m(m-1)ab(a+b)^{m-2} + m(m-1)(m-2)b^2a(a+b)^{m-3} + m(m-1)(m-2)ba^2(a+b)^{m-3} + \\ + \quad m(m-1)(m-2)(m-3)a^2b^2(a+b)^{m-4} = \sum_{j=0}^{m} j^2(m-j)^2 a^j b^{m-j} \binom{m}{j} \qquad (51)$$

Problem 3.9

- Now again take the partial derivative with respect to *b*

$$
\begin{aligned}
& m(m-1)a(a+b)^{m-2} + m(m-1)(m-2)ba(a+b)^{m-3} + m(m-1)(m-2)a^2(a+b)^{m-3} + \\
+ \quad & m(m-1)(m-2)(m-3)a^2b(a+b)^{m-4} = \sum_{j=0}^{m} j^2(m-j)^2 a^j b^{m-j-1} \binom{m}{j} \qquad (50)
\end{aligned}
$$

- … then multiply by *b* once more:

$$
\begin{aligned}
& m(m-1)ab(a+b)^{m-2} + m(m-1)(m-2)b^2a(a+b)^{m-3} + m(m-1)(m-2)ba^2(a+b)^{m-3} + \\
+ \quad & m(m-1)(m-2)(m-3)a^2b^2(a+b)^{m-4} = \sum_{j=0}^{m} j^2(m-j)^2 a^j b^{m-j} \binom{m}{j} \qquad (51)
\end{aligned}
$$

Problem 3.9

- Set $a = b = 1$:

$$\begin{aligned} & m(m-1)2^{m-2} + m(m-1)(m-2)2^{m-3} + m(m-1)(m-2)2^{m-3} + \\ + \quad & m(m-1)(m-2)(m-3)2^{m-4} = \sum_{j=0}^{m} j^2 (m-j)^2 \binom{m}{j} \end{aligned} \tag{52}$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Rearranging a bit:

$$
\begin{aligned}
&\sum_{j=0}^{m} j^2(m-j)^2\binom{m}{j} = m(m-1)2^{m-4}\left(4+2(m-2)+2(m-2)+(m-2)(m-3)\right) = \\
= \;& 2^{m-4}m(m-1)(m^2-m+2) \qquad (53)
\end{aligned}
$$

Problem 3.9

- Rearranging a bit:

$$
\begin{aligned}
\overline{F}_2 &= 2^{2-m}\frac{1}{m(m-1)}\sum_{j=0}^{m} j^2(m-j)^2\binom{m}{j} = \\
&= 2^{2-m}\frac{1}{m(m-1)}2^{m-4}m(m-1)(m^2-m+2) = \frac{m(m-1)}{4}+\frac{1}{2}
\end{aligned}
\tag{54}
$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Problem 3.9

- Rearranging a bit:

$$
\begin{aligned}
\overline{F}_2 &= 2^{2-m}\frac{1}{m(m-1)}\sum_{j=0}^{m} j^2(m-j)^2\binom{m}{j} = \\
&= 2^{2-m}\frac{1}{m(m-1)}2^{m-4}m(m-1)(m^2-m+2) = \frac{m(m-1)}{4}+\frac{1}{2}
\end{aligned}
\tag{54}
$$

Mattias Wahde, PhD, Professor, Chalmers University of Technology
e-mail: mattias.wahde@chalmers.se, http://www.me.chalmers.se/~mwahde

Today’s learning goals

- After this lecture you should be able to
  - Solve problems 2.12 and 2.13 in the book
  - Solve problems 3.2 and 3.9 in the book