Chalmers University of Technology, Department of Mechanics and Maritime Sciences Teacher visiting the exam room: Krister Blanch, Tel. 0700-277641

Exam in FFR105/FIM711 (Stochastic optimization algorithms), 2023-10-25, 08.30-12.30

The teacher will visit the exam room twice, around 09.30 and around 11.30. It will be possible to review your results (exam and home problems) from Nov. 17 and onwards.

In the exam, it is allowed to use an approved ("typgodkänd") calculator. No other calculators are allowed. Furthermore, it is *not* allowed to use the course book, mathematical tables, lecture notes, or slides from the course, during the exam.

Note! For Problem 1 below, you should *only* give the answers, i.e., a single letter (A), (B), (C), (D), or (E) for each question, all collected on a single page. For Problems 2 - 4 you should provide a clear description of how you arrived at your answer, including all relevant intermediate steps. *Only giving the answer will not give any points for Problems* 2 - 4.

There are four problems in the exam, and the maximum number of points is 25.

- This problem consists of ten questions, each of which is associated with five alternative answers, of which one is correct. For each of the 10 questions below, you should therefore only give one answer, in the form of a letter: (A), (B), (C), (D), or (E). Answer the questions in order (1.1, 1.2, ...,1.10) and mark your answers *clearly* (e.g., "1.1: A"). For this problem, the number of points awarded is determined as follows: 10 correct answers: 10p, 9 correct answers: 8p, 8 correct answers: 6p, 7 correct answers: 5p, 6 correct answers: 4p, 5 correct answers: 3p, 4 correct answers: 2p, 3 correct answers: 1p, 2 or fewer correct answers: 0p.
 - 1.1. Consider a binary chromosome in a GA, which has been generated via selection and crossover, and is then going to be mutated using the standard mutation procedure, with mutation rate $p_{mut} = 1/m$, where m is the chromosome length. Which of the following statements is true?
 - A. Either 0 or 1 gene will undergo mutation.
 - B. Exactly 1 gene will undergo mutation.
 - C. Either 0, 1, or 2 genes will undergo mutation.
 - D. The number of genes that undergo mutations will be in the range [0, m/2].
 - E. The number of genes that undergo mutation will be in the range [0, m].
 - 1.2. When using PSO, it is essential to keep the swarm coherent. Which of the following statements is true?
 - A. Only positions should be restricted.
 - B. Only velocities should be restricted.
 - C. Both velocities and positions should be restricted.
 - D. Positions should be restricted if particles venture outside the initial range.
 - E. Positions should be restricted if the inertia is smaller than 1.

- 1.3. Consider the function $f(x_1, x_2) = 4x_1^2 + 2x_2^2 2x_1x_2$ and let *H* denote the Hessian matrix. Which of the following statements is true?
 - A. The eigenvalues of H are both positive, so f is convex.
 - B. The eigenvalues of H are both positive, so f is not convex.
 - C. The eigenvalues of H are both negative, so f is convex.
 - D. The eigenvalues of H are both negative, so f is not convex.
 - E. One eigenvalue of H is positive, and one is negative, so f is not convex.
- 1.4. Which of the following statements is true? In a PSO run, the inertia term $(w) \dots$
 - A. ... should always be strictly smaller than 1.
 - B. ... should always be equal to 1.
 - C. ... should always be strictly larger than 1.
 - D. ... should be larger than 1 in the beginning, and then fall off to values smaller than 1 at the end of the run.
 - E. ... should be larger than 1 in the beginning, and then fall off until it reaches 1.
- 1.5. Consider a genetic algorithm in which individuals are selected using tournament selection, with a given value of p_{tour} and with a tournament size of three. In a single selection step (selecting one individual from the tournament with three individuals), what is the probability (p) of selecting the second-best individual.
 - A. $p = p_{tour}^2$ B. $p = p_{tour}^3$ C. $p = (1 - p_{tour})p_{tour}$
 - D. $p = (1 p_{tour})p_{tour}^2$
 - E. $p = (1 p_{tour})^2$
- 1.6. Consider a genetic algorithm with a population of five individuals, with fitness values $F_1 = 1$, $F_2 = 3$, $F_3 = 5$, $F_4 = 8$, and $F_5 = 10$. Assuming that tournament selection is used, with $p_{tour} = 0.75$, what is the probability (in a single selection step) of selecting individual 2 (whose fitness is equal to 3)?
 - A. 0.10
 - B. 0.12
 - C. 0.16
 - D. 0.18
 - E. 0.24
- 1.7. Consider the ant system (AS) algorithm. Let τ_{ij} denote the pheromone matrix and η_{ij} the visibility matrix. For any problem where AS is applied, at the end of a run lasting a given (finite) number of iterations, ...
 - A. ... the pheromone matrix is symmetric, whereas the visibility matrix is not.
 - B. ... the visibility matrix is symmetric, whereas the pheromone matrix is not.
 - C. ... both the visibility matrix and the pheromone matrix are symmetric.
 - D. ... neither the visibility matrix nor the pheromone matrix is symmetric.
 - E. ... the symmetry (or lack thereof) of either matrix depends on the problem.

- 1.8. Consider a case in which the Lagrange multiplier method is applied to the problem of finding the minima of a continuously differentiable function $f(x_1, x_2)$ subject to a continuously differentiable equality constraint $h(x_1, x_2) = 0$, both defined in a bounded region (i.e., such that it can be contained in a disc of radius R, where R is finite). Which of the following statements is true? For the problem at hand, the points found by the Lagrange multiplier method ...
 - A. ... contain local and global minima, but no maximum, neither local nor global.
 - B. ... contain local and global maxima, but no minimum, neither local nor global.
 - C. ... contain local minima and maxima, but not any global minima or maxima.
 - D. ... contain only global minima and maxima.
 - E. ... contain local and global optima, both minima and maxima.
- 1.9. Consider a case in which ant system (AS) is being used for solving the travelling salesperson problem (TSP), and where three nodes remain to be selected to complete the path of a given ant. From the current node, the visibilities of those nodes are 1, 2, and 5, respectively. Moreover, the parameters α and β are both equal to 1, and the pheromone level (τ) is equal to 1 on all edges. What is the probability of selecting the third node (with visibility 5) as the next node?
 - A. 0
 - B. 1/3
 - C. 3/8
 - D. 5/8
 - E. 1
- 1.10. Consider a standard genetic algorithm (GA) where either tournament selection (TS) or (standard) roulette-wheel selection (RWS), without fitness ranking, is used, and where the fitness values (f) are in the range [0,1]. Which of the following statements is true? An individual whose fitness is equal to 0 ...
 - A. ... can be selected with RWS but not with TS.
 - B. ... can be selected with TS but not with RWS.
 - C. ... cannot be selected with either method.
 - D. ... can be selected with both methods.
 - E. ... can be selected with both methods, but only in the first generation.

2) In gradient descent one follows the negative gradient from a given starting point towards a (local) minimum. In this method, starting from a given point in x_j (where the subscript in this case enumerates the iterations, starting from j = 0) n dimensions, one finds iterates such that, once the search direction has been determined, the next iterate x_{j+1} will depend only on the step length η , so that the function value at that point can be expressed as some function $\Phi(\eta)$. Consider now the problem of finding the minimum of the function

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

(where, in this case, the indices of course enumerate the components of x) using gradient descent, starting from the point $x = (x_1, x_2) = (1, 1)$.

- (a) Find the search direction (i.e., a vector with two components) at (1,1), as well as the expression for the next iterate, inserting numerical values. (1p)
- (b) Derive the expression for the polynomial $\Phi(\eta)$. (1p)
- (c) Using your calculator sketch (and include in your solution) $\Phi(\eta)$ (or an equivalent function, if you wish to make a variable substitution based on your findings in (b)), so that you get a rough idea of where the minimum (or minima) are located. (1p)
- (d) Next, carry out a line search using Newton-Raphson's method, in order to find the location η^* of the (global) minimum of $\Phi(\eta)$, with at least 5-decimal precision (2p).
- (e) Given the value of η^* from (d), find the corresponding point (x_1^*, x_2^*) , and then determine the gradient at this point. Next, compute the scalar product between this gradient and the gradient from (a), and briefly discuss your findings. (1p)

3	2	1	6	3	3	2	2	1	2	1	5	1	3	2	3	3	1	1	5	2	2	1	3



3) Consider the LGP chromosome shown in Figure 1. As usual in LGP, the chromosome defines a list of instructions, each consisting of four genes. For each instruction, the first gene defines the operator from the set 0 = {+, -, *, /} (i.e., the four standard arithmetic operators), such that 1 ⇔ +, 2 ⇔ -, and so on. The second gene defines the destination register, from the set R = {r1, r2, r3}. The third and fourth genes define the operands, taken from the set A = {r1, r2, r3, c1, c2, c3}. The constant registers take the values c1 = 1, c2 = 2, and c3 = -1. When evaluating the chromosome for a general variable x, the variable registers are initialized as follows: r1 = x, r2 = r3 = 0, and the output, namely a function g(x), is taken as the content of r2 at the end of the evaluation. What function g(x) is represented by the chromosome in Fig. 1? Write your answer in the form of a polynomial, g(x) = a₀ + a₁x + a₂x² + ... (3p)



- 4) In Figure 2, a construction graph for the travelling salesman problem (TSP) is shown. In this case, the aim is to find the shortest path, using ant system (AS). In the graph, there are three nodes (nodes 1,2, and 3) with fixed locations: Node 1 (denoted v_1 in the figure) is at (0,0), Node 2 at (0,2) and Node 3 at (0,4), whereas Node 4 is placed somewhere on the line segment from (1,0) to (1,4). This line segment (like any line segment) can be expressed as a function of a single parameter p, where p is in the range [0,1], in this case as $(x_1, x_2) = (1,4p)$.
 - as (x₁, x₂) = (1,4p).
 (a) Assuming that the ant starts at Node 1, find the sequence of nodes in the nearest-neighbor (NN) path, remembering that the sequence depends on the value of p. (In
 - neighbor (NN) path, remembering that the sequence depends on the value of p. (In cases where there are two options with equal length, consider both options.) (2p)
 - (b) Pheromone updates are made as usual in TSP, but with one difference: The amount of deposited pheromone is the inverse of what we here denote D^2 . The value of D^2 is computed by summing, over the entire path, the *square* of the length of each segment along the path (rather than just using the path length, as would normally be the case for TSP). Assuming that the ant starts at Node 1 and follows the NN path(s) from part (a), for what values of p will the deposited amount of pheromone be maximal, and what is that value? (4p)

Note: For part (b), points are only given if the results from part (a) are correct.

Stochastic optimization methods (FFR 105), 2023 Solutions to the exam (2023-10-25)

- 1. 1.1 E. Any number of genes (in the range [0, m]) can mutate, in principle, since the mutation is applied on a gene-by-gene basis.
 - 1.2 B. Only velocities should be restricted.
 - 1.3 A. Both eigenvalues are positive, so the function is convex.
 - 1.4 D. w starts at values larger than 1, usually around 1.4, and then falls off to values smaller than one, typically ending around 0.3-0.4.
 - 1.5 C. The probability equals $p = (1 p_{tour})p_{tour}$.
 - 1.6 C. The probability equals $p_2 = \frac{1}{25} \left(2 \times \frac{3}{4} + 6 \times \frac{1}{4} + 1 \right) = 0.16.$
 - 1.7 E. No general statement can be made regarding the symmetry of either matrix.
 - 1.8 E. The list of points found by applying the Lagrange multiplier method contains all optima of f subject to the constraint, both local and global ones.
 - 1.9 D. Since the pheromone levels are the same on all edges, the pheromone need not be considered. With $\beta = 1$, the probability of selecting node 3 equals

$$p = \frac{5}{1+2+5} = \frac{5}{8} \tag{1}$$

- 1.10 B. An individual with fitness 0 will not be selected with RWS, in which selection is directly proportional to the fitness values. However, it can be selected with TS, which only considers the fitness values (of the participants in the tournament) relative to each other.
- 2. (a) The first search direction (\mathbf{d}_0) is given by the negative gradient, which in this case takes the form

$$\mathbf{d}_0 = -\nabla f|_{\mathbf{X}=\mathbf{X}_0} = -(4x_1^3 - 2, -2x_2 + 4x_2^3)^{\mathrm{T}}.$$
 (2)

Inserting numerical values, one finds

$$\mathbf{d}_0 = -(2,2)^{\mathrm{T}}.$$
 (3)

The expression for the next iterate becomes

$$\mathbf{x}_1 = \mathbf{x}_0 - \eta \nabla f |_{\mathbf{X} = \mathbf{X}_0} = (1 - 2\eta, 1 - 2\eta)^{\mathrm{T}}.$$
(4)

(b) Using the fact that $\phi(\eta) \equiv f(\mathbf{x}_1(\eta))$, one finds

$$\phi(\eta) = (1 - 2\eta)^4 - 2(1 - 2\eta) - (1 - 2\eta)^2 + (1 - 2\eta)^4$$

= 2(1 - 2\eta)^4 - (1 - 2\eta)^2 - 2(1 - 2\eta). (5)

(c) As a simplification, set $p = 1 - 2\eta$, so that the function $\phi(\eta)$ can be written (equivalently) as $\nu(p) \equiv \phi(p(\eta)) = 2p^4 - p^2 - 2p$. Plotting $\nu(p)$ one obtains a curve as in Fig. 1 As is evident from the functional form, ν will rise very quickly



Figure 1: The function $\nu(p)$ from Problem 2.

for large values of |p|. By inspection, one finds that the global minimum will be somewhere between 0.5 and 1. Applying Newton-Raphson's method, one obtains new iterates as

$$p \leftarrow p - \frac{f'(p)}{f''(p)} \equiv p - \frac{8p^3 - 2p - 2}{24p^2 - 2}.$$
 (6)

Starting at p = 1 and iterating with this equation, one obtains the sequence $\{0.8181818, 0.7651960, 0.7607207, 0.7606899, 0.7606899, \ldots\}$. Thus, the method converges to $p^* \approx 0.7606899 \Rightarrow \eta^* = (1 - p^*)/2 = 0.1196551$.

(d) The corresponding point \mathbf{x}_1 is given (see (a) above) in component form as $(x_1^*, x_2^*) = (1 - 2\eta^*, 1 - 2\eta^*)^{\mathrm{T}} = (0.7606899, 07606899)^{\mathrm{T}}$. At this point, by inserting the values of x_1 and x_2 just given, the gradient (again, see (a) above) becomes

$$\nabla f|_{\mathbf{X}=\mathbf{X}_1} \approx (-0.23931, 0.23931)^{\mathrm{T}}$$
 (7)

Thus, the scalar product of this gradient and the gradient computed in (a) is zero. This is to be expected since, in gradient descent, successive search directions are orthogonal to each other (provided that one, in each step, finds the global minimum of $\phi(\eta)$, and then computes the next search direction at the point thus found).

- 3. Given the specification in the problem formulation, the instructions (here denoted $I_j, j = 1, ...6$) can be decoded as
 - $\begin{array}{rrrr} I_1 & 3216; & r_2 = r_1 \times c_3 \\ I_2 & 3322; & r_3 = r_2 \times r_2 \\ I_3 & 1215; & r_2 = r_1 + c_2 \\ I_4 & 1323; & r_3 = r_2 + r_3 \\ I_5 & 3115; & r_1 = r_1 \times c_2 \\ I_6 & 2213; & r_2 = r_1 r_3 \end{array}$

Starting from $r_1 = x$, $r_2 = r_3 = 0$, one then gets

	r_1	r_2	r_3
I_1	x	-x	0
I_2	x	-x	x^2
I_3	x	x+2	x^2
I_4	x	x+2	$x^2 + x + 2$
I_5	2x	x+2	$x^2 + x + 2$
I_6	2x	$-x^2 + x - 2$	$x^2 + x + 2$

Hence, the function obtained (from r_2) at the end of the calculation is

$$g(x) = -x^2 + x - 2.$$
(8)

4. (a) Starting at Node 1 (as was given), the nearest node is either Node 2 or Node 4, depending on the value of p. For p = 0, Node 4 is at a distance of 1 from Node 1, which is smaller than the distance (=2) from Node 1 to Node 2. We note that d²₂₁ = 4, whereas

$$d_{41}^2 = 1^2 + (4p)^2 = 16p^2 + 1.$$
(9)

Thus, the transistion point where Node 2 and Node 4 are at the same distance from Node 1 occurs when

$$16p^2 + 1 = 4 \Rightarrow p = \sqrt{\frac{3}{16}}.$$
 (10)

For $0 \le p \le \sqrt{3/16}$ we thus find that the NN path takes the form $(\nu_1, \nu_4, \nu_2, \nu_3)$, since Node 4 is clearly closer to Node 2 than Node 3. For $p \ge \sqrt{3/16}$, the start of the sequence is (ν_1, ν_2) (for the case $p = \sqrt{3/16}$ two paths are possible, so we can list both paths as possibilities). The identity of the third node again depends on the value of p. The relevant squared distances are $d_{32}^2 = 4$ and

$$d_{42}^2 = 1^2 + (4p - 2)^2 = 16p^2 - 16p + 5.$$
(11)

Thus, the value of p where Node 3 and Node 4 are at an equal distance from Node 2 occurs when

$$16p^2 - 16p + 5 = 4 \Rightarrow 16p^2 - 16p + 1 = 0 \Rightarrow p^2 - p + \frac{1}{16} = 0, \quad (12)$$

for which one finds the solutions

$$p = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{16}} = \frac{1}{2} \pm \sqrt{\frac{3}{16}}.$$
 (13)

Here it is only the solution $p = 1/2 + \sqrt{3/16}$ that is relevant. This is easy to see: For p < 1/2, Node 4 (rather than Node 3) is clearly closer to Node 2. Thus, we find that the NN path takes the form $(\nu_1, \nu_2, \nu_4, \nu_3)$ for $\sqrt{3/16} \le p \le 1/2 + \sqrt{3/16}$, and the form $(\nu_1, \nu_2, \nu_3, \nu_4)$ for $1/2 + \sqrt{3/16} \le p \le 1$. At the transition point, $p = 1/2 + \sqrt{3/16}$ we should consider both possibilities.

(b) The amount of deposited pheromone is equal to $1/D^2$ where D^2 is the sum of the lengths of each segment in the path traversed by the ant. Here, there are three cases:

i. $0 \le p \le \sqrt{\frac{3}{16}}$: In this case, D^2 is given by $D_1^2 = d_{41}^2 + d_{24}^2 + d_{32}^2 + d_{13}^2 =$ $= 16p^2 + 1 + 16p^2 - 16p + 5 + 4 + 16 = 32p^2 - 16p + 26$ (14)

ii. $\sqrt{\frac{3}{16}} \le p \le 1/2 + \sqrt{3/16}$: Here, D^2 is given by $D_2^2 = d_{21}^2 + d_{42}^2 + d_{34}^2 + d_{13}^2 = 4 + 16p^2 - 16p + 5 + 16p^2 - 32p + 17 + 16 = 32p^2 - 48p + 42$ (15)

since

$$d_{34}^2 = 1^2 + (4p - 4)^2 = 16p^2 - 32p + 17$$
(16)

iii. $1/2 + \sqrt{\frac{3}{16}} \le p \le 1$: In this case, D^2 is computed as

$$D_3^2 = d_{21}^2 + d_{32}^2 + d_{43}^2 + d_{14}^2 = = 4 + 4 + 16p^2 - 32p + 17 + 16p^2 + 1 = 32p^2 - 32p + 26.$$
(17)

In order to maximize $1/D^2$ one can minimize D^2 . In all three cases above, the squared NN path length is parabola, with positive coefficient for the quadratic term. The minima can occur either at p = 0 or p = 1, or at the boundaries between the three cases, or at the (single) stationary point for each case. We need to include the edges for each case since, as clearly stated in the problem formulation, in cases where there are two options with equal length, consider both options.

However, if, for a given case, the stationary point falls within the interval, then there is no need to consider the edges, since the global minimum of the (convex) parabola is at the stationary point. Thus, in theory, we have a total of $3 \times 3 = 9$ cases to consider, but we start by investigating the stationary points for each case.

Case i: Set $d(D_1^2)/dp = 0$. From the equation for D_1^2 we have $d(D_1^2)/dp = 64p - 16$, so that p = 1/4. We note that $1/4 < \sqrt{3/16}$. Thus, the stationary point is indeed in the interval defining D_1 , so we need not consider the edges $(p = 0 \text{ and } p \rightarrow \sqrt{3/16})$. Inserting this stationary point we find $D_1^2 = 32/16 - 16/4 + 26 = 24$.

Case ii: Set $d(D_2^2)/dp = 0$. From the equation for D_2^2 we have $d(D_2^2)/dp = 64p - 48$, so that p = 3/4. We note that $\sqrt{3/16} < 3/4 < 1/2 + \sqrt{3/16}$ so that the stationary point is in the interval defining D_2 , meaning that here, too, we need only consider the stationary point. Inserting the point found in the expression for D_2^2 , we get $D_2^2 = 32 \times 9/16 - 48 \times 3/4 + 42 = 24$.

Case iii: Set $d(D_3^2)/dp = 0$. From the equation for D_3^2 we have $d(D_3^2)/dp = 64p - 32$, so that p = 1/2. However, this point is clearly not in the interval defining D_3 , so it need not be considered further. Instead we need to consider the two edges. First, set $p = 1/2 + \sqrt{3/16}$. Here we get $D_3^2 = 32 \times (1/2 + \sqrt{3/16})^2 - 32(1/2 + \sqrt{3/16}) + 26 = 24$. Next, for p = 1, we get $D_3^2 = 32 - 32 + 26 = 26$.

Thus, looking at all the cases above, we find that D^2 is minimal for p = 1/4, p = 3/4, and $p = 1/2 + \sqrt{3/16}$ (approaching from above, so that the NN path is $(\nu_1, \nu_2, \nu_3, \nu_4)$) where it takes the value 24. The deposited amount of pheromone is equal to $1/D^2 = 0.041667$.