

**Federal State Autonomous Educational Institution of Higher Education "Moscow  
Institute of Physics and Technology  
(National Research University)"**

**APPROVED**

**Head of the Phystech School of  
Applied Mathematics and  
Informatics**

**A.M. Raygorodskiy**

**Work program of the course (training module)**

<b>course:</b>	Optimization Methods/Методы оптимизации
<b>major:</b>	Applied Mathematics and Informatics
<b>specialization:</b>	Modern State of Artificial Intelligence/Современные методы искусственного интеллекта “Pusk” Online and Supplementary Education Centre Chair of Machine Learning and Digital Humanities
<b>term:</b>	1
<b>qualification:</b>	Master

Semester, form of interim assessment: 1 (fall) - Exam

Academic hours: 60 AH in total, including:

lectures: 30 AH.

seminars: 30 AH.

laboratory practical: 0 AH.

Independent work: 45 AH.

Exam preparation: 30 AH.

In total: 135 AH, credits in total: 3

Author of the program: A.V. Gasnikov, doctor of physics and mathematical sciences, associate professor, professor

The program was discussed at the Chair of Machine Learning and Digital Humanities 05.03.2020

## Annotation

This course presents an overview of the methods and techniques from the optimization that are crucial in understanding machine learning and deep learning methods. We start with the classical convexity notion, review the main facts from the convex optimization, highlight the most practically important numerical methods, and finally consider relaxations of the discrete optimization problems. The home assignments include both theoretical questions and programming exercises.

### 1. Study objective

#### Purpose of the course

- Learn the main theoretical foundations of widely used optimization methods and their limitations
- Gain essential experience in formulating, decomposing and solving various optimization problems
- Get familiar with various approaches optimization problems

#### Tasks of the course

1. Solve medium-scale convex optimization problems with modern solvers
2. Convexity verification of the given optimization problem
3. Introducing stochasticity in the optimization methods
4. Understanding the pros and cons of the standard large-scale optimization methods

### 2. List of the planned results of the course (training module), correlated with the planned results of the mastering the educational program

Mastering the discipline is aimed at the formation of the following competencies:

Code and the name of the competence	Competency indicators
Gen.Pro.C-1 Address current challenges in fundamental and applied mathematics	Gen.Pro.C-1.2 Consolidate and critically assess professional experience and research findings
Pro.C-1 Become part of a professional community and conduct local research under scientific guidance using methods specific to a particular professional setting	Pro.C-1.2 Understand the verification process of software models used to solve related scientific problems
Pro.C-2 Understands and is able to apply modern mathematical apparatus and algorithms, the basic laws of natural science, modern programming languages and software; operating systems and networking technologies in research and applied activities	Pro.C-2.1 Demonstrate expert knowledge of research basics in the field of ICTs, philosophy and methodology of science, scientific research methods, and apply skills to use them

### 3. List of the planned results of the course (training module)

As a result of studying the course the student should:

know:

- necessary and sufficient conditions for the extrema of functions of several variables;
- methods of searching for extrema;
- the structure of the dynamic system;
- methods for finding optimal control.

be able to:

- to reduce the optimization problem to the canonical form;
- choose an optimization method based on the characteristics of the problem;
- build an algorithm for finding the optimal control

master:

- methods for finding the extrema of functions of many variables;
- numerical optimization approaches;
- the basics of the calculus of variations.

#### 4. Content of the course (training module), structured by topics (sections), indicating the number of allocated academic hours and types of training sessions

##### 4.1. The sections of the course (training module) and the complexity of the types of training sessions

№	Topic (section) of the course	Types of training sessions, including independent work			
		Lectures	Seminars	Laboratory practical	Independent work
1	Convex sets and convex functions	6	6		9
2	Optimality conditions	6	6		9
3	Intro to numerical methods, gradient descent	6	6		9
4	Quasi-Newton optimization methods	6	6		9
5	Intro to discrete problems	6	6		9
AH in total		30	30		45
Exam preparation		30 AH.			
Total complexity		135 AH., credits in total 3			

##### 4.2. Content of the course (training module), structured by topics (sections)

Semester: 1 (Fall)

###### 1. Convex sets and convex functions

Matrix calculus, automatic differentiation    Subgradient and subdifferential

###### 2. Optimality conditions

Duality and conjugate things    Modeling in convex optimization

###### 3. Intro to numerical methods, gradient descent

Accelerated modifications of gradient descent

###### 4. Quasi-Newton optimization methods

SGD and its modifications in theory and in practice

###### 5. Intro to discrete problems

Convex relaxations of combinatorial optimization problems

#### 5. Description of the material and technical facilities that are necessary for the implementation of the educational process of the course (training module)

A standard classroom.

## **6. List of the main and additional literature, that is necessary for the course (training module) mastering**

### Main literature

1. Методы оптимизации. Условия оптимальности в экстремальных задачах [Текст] : учеб. пособие для вузов / А. Г. Бирюков ; М-во образования и науки Рос. Федерации, Моск. физ.-техн. ин-т (гос. ун-т) .— М. : МФТИ, 2010 .— 225 с.
2. Методы оптимизации [Текст]. Ч. 2. Численные алгоритмы : учеб. пособие для вузов / Жадан, В. Г. ; М-во образования и науки РФ, Моск. физ.-техн. ин-т (гос. ун-т) .— М. : МФТИ, 2015 .— 320 с. + pdf-версия. - Библиогр.: с. 314-319. - 300 экз. - ISBN 978-5-7417-0571-1. — Полный текст (Доступ из сети МФТИ).

### Additional literature

1. Курс методов оптимизации [Текст] : учеб. пособие для вузов / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров ; [Моск. гос. ун-т им. М. В. Ломоносова] .— 2-е изд. — М. : Физматлит, 2005, 2008 .— 367 с.

## **7. List of web resources that are necessary for the course (training module) mastering**

<http://dm.fizteh.ru/>

## **8. List of information technologies used for implementation of the educational process, including a list of software and information reference systems (if necessary)**

Multimedia technologies can be employed during lectures and practical lessons, including presentations.

## **9. Guidelines for students to master the course**

A student studying a discipline must, on the one hand, master the general conceptual apparatus, and on the other hand, must learn to apply theoretical knowledge in practice.

As a result of studying the discipline, the student must know the basic definitions and concepts.

Successful mastering of the course requires intense independent work of the student. The course program provides the minimum required time for a student to work on a topic. Independent work includes:

- reading and taking notes of the recommended literature;
- study of educational material (based on lecture notes, educational and scientific literature);
- preparation for differential credit and exam.

Guidance and control over the student's independent work is carried out in the form of individual consultations.

It is important to achieve an understanding of the studied material, and not its mechanical memorization. If you find it difficult to study certain topics, questions, you should seek advice from a teacher.

**Assessment funds for course (training module)**

**major:** Applied Mathematics and Informatics  
**specialization:** Modern State of Artificial Intelligence/Современные методы искусственного интеллекта  
“Pusk” Online and Supplementary Education Centre  
Chair of Machine Learning and Digital Humanities  
**term:** 1  
**qualification:** Master

Semester, form of interim assessment: 1 (fall) - Exam

**Author:** A.V. Gasnikov, doctor of physics and mathematical sciences, associate professor, professor

## 1. Competencies formed during the process of studying the course

Code and the name of the competence	Competency indicators
Gen.Pro.C-1 Address current challenges in fundamental and applied mathematics	Gen.Pro.C-1.2 Consolidate and critically assess professional experience and research findings
Pro.C-1 Become part of a professional community and conduct local research under scientific guidance using methods specific to a particular professional setting	Pro.C-1.2 Understand the verification process of software models used to solve related scientific problems
Pro.C-2 Understands and is able to apply modern mathematical apparatus and algorithms, the basic laws of natural science, modern programming languages and software; operating systems and networking technologies in research and applied activities	Pro.C-2.1 Demonstrate expert knowledge of research basics in the field of ICTs, philosophy and methodology of science, scientific research methods, and apply skills to use them

## 2. Competency assessment indicators

As a result of studying the course the student should:

### know:

- necessary and sufficient conditions for the extrema of functions of several variables;
- methods of searching for extrema;
- the structure of the dynamic system;
- methods for finding optimal control.

### be able to:

- to reduce the optimization problem to the canonical form;
- choose an optimization method based on the characteristics of the problem;
- build an algorithm for finding the optimal control

### master:

- methods for finding the extrema of functions of many variables;
- numerical optimization approaches;
- the basics of the calculus of variations.

## 3. List of typical control tasks used to evaluate knowledge and skills

1. Examples of applications that can be reduced to convex optimization problems
2. Definitions of a convex set, convex function, and epigraph
3. KKT optimality conditions, their sufficiency and necessity
4. What benefits can give a dual problem?
5. Convergence of the gradient descent method and approaches to accelerate it
6. Comparison of the gradient descent and Newton methods
7. What is the computational complexity of the L-BFGS method? Why?
8. What problems are appropriate to be solved with SGD? What is the convergence of SGD and how it can be accelerated?
9. How to construct convex relaxation for the nonconvex quadratic programming problems?

## 4. Evaluation criteria

Questions for the exam

1. Prove that if  $m, n$  are two coprime integers of different parity, then the numbers  $m^2 - n^2$  and  $2mn$  are also coprime.

2. Write and prove the general formula for the number of different representations of a given integer  $n$  as the sum of two squares. (Representatives that are not obtained from each other by changing signs and the order of the terms are considered different.)
3. Based on the obtained formula, derive the lower bound for the maximum number of equal distances among the given  $n$  points on the plane using a regular rectangular lattice.
4. Build a regular pentagon using a compass and a ruler.
5. Build a regular 15-gon using a compass and a ruler.
6. You are given a single segment. It is required to construct using a compass and a ruler a segment of length  $x$  satisfying the equation
7. Based on the previous task, prove that a regular heptagon cannot be built using a compass and a ruler.
8. Prove that trisection of the angle is impossible.
9. Describe all possible combinations of the amounts of black and white balls in the ballot box, so that if two balls are randomly fished in a sample without returning, the probability of fishing two white balls is exactly 0.5.
10. Consider the relation on the sides  $a$ ,  $b$ ,  $c$  of the triangle, in which a triangle with vertices at the bases of the bisectors is isosceles. Assuming that the sides converging on side  $c$  of the large triangle are equal, reduce this relation to the following
11. In what follows, we consider the cube defined by the first of the three equations (refusing the requirement that  $a$ ,  $b$ ,  $c$  be sides of a triangle). Show that the resulting cube is indecomposable, that is, the polynomial that defines it does not factor.
12. In addition to this, show that our cube is nonsingular, that is, there is not a single point on its projectivization at which each direction is tangent (or the same thing at which all three first partial derivatives of the polynomial defining it degenerate).

#### Exam ticket examples

##### Ticket number 1

1. Write and prove the general formula for the number of different representations of a given integer  $n$  as the sum of two squares.
2. Prove that trisection of the angle is impossible.

##### Ticket number 2

1. Consider the relationship on the sides  $a$ ,  $b$ ,  $c$  of the triangle, in which a triangle with vertices at the bases of the bisectors is isosceles.
2. Describe all kinds of combinations of the numbers of black and white balls in the ballot box, so that if two balls are randomly fished in the sample without returning, the probability of fishing two white balls is exactly 0.5.

Assessment “excellent (10)” is given to a student who has displayed comprehensive, systematic and deep knowledge of the educational program material, has independently performed all the tasks stipulated by the program, has deeply studied the basic and additional literature recommended by the program, has been actively working in the classroom, and understands the basic scientific concepts on studied discipline, who showed creativity and scientific approach in understanding and presenting educational program material, whose answer is characterized by using rich and adequate terms, and by the consistent and logical presentation of the material;

Assessment “excellent (9)” is given to a student who has displayed comprehensive, systematic knowledge of the educational program material, has independently performed all the tasks provided by the program, has deeply mastered the basic literature and is familiar with the additional literature recommended by the program, has been actively working in the classroom, has shown the systematic nature of knowledge on discipline sufficient for further study, as well as the ability to amplify it on one’s own, whose answer is distinguished by the accuracy of the terms used, and the presentation of the material in it is consistent and logical;

Assessment “excellent (8)” is given to a student who has displayed complete knowledge of the educational program material, does not allow significant inaccuracies in his answer, has independently performed all the tasks stipulated by the program, studied the basic literature recommended by the program, worked actively in the classroom, showed systematic character of his knowledge of the discipline, which is sufficient for further study, as well as the ability to amplify it on his own;

Assessment “good (7)” is given to a student who has displayed a sufficiently complete knowledge of the educational program material, does not allow significant inaccuracies in the answer, has independently performed all the tasks provided by the program, studied the basic literature recommended by the program, worked actively in the classroom, showed systematic character of his knowledge of the discipline, which is sufficient for further study, as well as the ability to amplify it on his own;

Assessment “good (6)” is given to a student who has displayed a sufficiently complete knowledge of the educational program material, does not allow significant inaccuracies in his answer, has independently carried out the main tasks stipulated by the program, studied the basic literature recommended by the program, showed systematic character of his knowledge of the discipline, which is sufficient for further study;

Assessment “good (5)” is given to a student who has displayed knowledge of the basic educational program material in the amount necessary for further study and future work in the profession, who while not being sufficiently active in the classroom, has nevertheless independently carried out the main tasks stipulated by the program, mastered the basic literature recommended by the program, made some errors in their implementation and in his answer during the test, but has the necessary knowledge for correcting these errors by himself;

Assessment “satisfactory (4)” is given to a student who has discovered knowledge of the basic educational program material in the amount necessary for further study and future work in the profession, who while not being sufficiently active in the classroom, has nevertheless independently carried out the main tasks stipulated by the program, learned the main literature but allowed some errors in their implementation and in his answer during the test, but has the necessary knowledge for correcting these errors under the guidance of a teacher;

Assessment “satisfactory (3)” is given to a student who has displayed knowledge of the basic educational program material in the amount necessary for further study and future work in the profession, not showed activity in the classroom, independently fulfilled the main tasks envisaged by the program, but allowed errors in their implementation and in the answer during the test, but possessing necessary knowledge for elimination under the guidance of the teacher of the most essential errors;

Assessment “unsatisfactory (2)” is given to a student who showed gaps in knowledge or lack of knowledge on a significant part of the basic educational program material, who has not performed independently the main tasks demanded by the program, made fundamental errors in the fulfillment of the tasks stipulated by the program, who is not able to continue his studies or start professional activities without additional training in the discipline in question;

Assessment “unsatisfactory (1)” is given to a student when there is no answer (refusal to answer), or when the submitted answer does not correspond at all to the essence of the questions contained in the task.

## **5. Methodological materials defining the procedures for the assessment of knowledge, skills, abilities and/or experience**

During examination the student are allowed to use the program of the discipline.