Львівський державний університет безпеки життєдіяльності ДСНС України

Кафедра прикладної математики і механіки

ТЕМА 5. НЕЛІНІЙНЕ ПРОГРАМУВАННЯ.
ДЛЯ ПРОВЕДЕННЯ ЛЕКЦІЙНОГО ЗАНЯТТЯ
З КУРСАНТАМИ ТА СТУДЕНТАМИ 2 КУРСУ
З ДИСЦИПЛІНИ СИСТЕМНИЙ АНАЛІЗ ТА ТЕОРІЯ ПРИЙНЯТТЯ РІШЕНЬ

Мета лекції:
Навчальна: вивчити поняття нелінійного програмування, метод множників Лагранжа, метод поділу відрізка навпіл.
Виховна: виховання свідомого ставлення до вивчення предмету, самостійності, відповідальності та організованості при підготовці до занять.
Розвиткова: розвиток логічного та абстрактного мислення, розвиток просторової уяви.

План

Література

1. *Аршинова О.І., Шевченко А.В.* Системний аналіз. Навч. посібник. – К.: НАУ, 2008. – 128 с.
2. *Роїк О.М., Шиян А.А., Нікіфорова Л.О.* Системний аналіз. Навч. посібник. – Вінниця: ВНТУ, 2015. – 83 с.
3. *Кузик А.Д.* Основи системного аналізу. – Львів: ЛДУ БЖД, 2005. – 100 с.
4. *Кунда Н.Т.* Дослідження операцій у транспортних системах. – К.: Видавничий Дім “Слово”, 2008. – 400 с.
5. *Карагодова О.О., Кігель В.Р., Рожок В.Д.* Дослідження операцій. – К.: Центр учбової літератури, 2007.– 256 с.
6. *Таха Х.* Введение в исследование операций. –6-е изд.: Пер. с англ. –М.: Изд. дом "Вильямс", 2001. – 912 с.

Час проведення: 2 години.
Місце проведення: лекційний зал.
Забезпечення заняття: мультимедіа.

1. Поняття нелінійного програмування.

Задачею нелінійного програмування називають таку задачу математичного програмування, у якій функція мети та/або система обмежень є нелінійними.

Виникають задачі нелінійного програмування в практичних завданнях досить часто. Наприклад, нехай критерієм оптимальності є собівартість одиниці виробленої продукції. Очевидно, що вона залежить від розміру підприємства. Так, із збільшенням обсягу продукції собівартість її зменшується. Проте таке зменшення не безмежне. Настає такий момент, коли внутрішні витрати підприємства починають зростати (збільшуються витрати на перевезення, збереження продукції тощо), що у свою чергу призводить до збільшення собівартості. Функція, яка і спадає, і зростає, вже не може бути лінійною. Крім того, якщо врахувати в моделях лінійного програмування додаткові фактори, то ці моделі трансформуються в нелінійні. Наприклад, припустивши, що в задачі про використання ресурсів обсяг реалізації впливає на прибуток, маємо функцію мети з нелінійністю.

Класифікують задачі нелінійного програмування <u>за наявністю обмежень</u> (задачі умовної та безумовної оптимізації) та <u>за кількістю змінних</u> (однопараметричні та багатопараметричні задачі оптимізації).

На відміну від лінійного програмування не існує загальних методів розв'язання задач нелінійного програмування. В кожному конкретному випадку метод вибирається з урахуванням виду функції мети (дробово-лінійна, квадратична та ін.). У випадку 2-х змінних для розв'язання задач нелінійного програмування можна використовувати графічний метод, аналогічно як до задач лінійного програмування. Деякі задачі нелінійного програмування можна наблизити до задач лінійного програмування та знайти розв'язок близький до оптимального.

Методи розв'язання задач нелінійного програмування знаходяться у процесі розвитку, тому їх використання є обмеженим. Найчіткіше розроблено методи розв'язування нелінійних задач з опуклою областю допустимих розв'язків та опуклою чи увігнутою функцією мети. Такі задачі складають клас задач опуклого програмування і є частковим випадком задач нелінійного програмування. Методи розв'язання задач опуклого програмування можна умовно поділити на дві групи. До першої групи належать методи, які зводять задачу до задачі безумовної оптимізації. У цьому випадку розв'язок задачі знаходиться як межа послідовності безумовних екстремумів за допомогою спеціально складених допоміжних функцій. До них включають метод множників Лагранжа, метод штрафних функцій, метод центрів.

Для знаходження стаціонарної точки цими методами треба розв'язати систему n нелінійних рівнянь з n змінними. Проте розв'язування такої системи є складною самостійною задачею. Крім того, задачі, що зустрічаються на практиці, мають цільову функцію не завжди диференційовну, що також утруднює процес розв'язання задачі.

До другої групи входять методи, що базуються на ідеї поступового наближення до точки екстремуму. При цьому має виконуватися умова монотонної зміни функції мети. Такі методи розв'язування нелінійних задач називають *ітеративними*. Основним класом цієї групи є градієнтні методи: метод покоординатного спуску (метод Гаусса-Зейделя), метод крутого спуску (метод Бокса-Уілсона) або метод найшвидшого спуску та інші.

У випадку, коли для задач нелінійного програмування не вдається побудувати ефективний алгоритм розв'язання, то використовують деякі штучні підходи:

- зведення нелінійної задачі до задачі лінійного програмування, якщо це можливо, не впливаючи на точність і зміст розв'язку; це має місце у випадках слабкої нелінійної залежності;
- зведення нелінійної задачі до структури багатокрокових задач, що дозволяє використовувати метод динамічного програмування.

Розглянемо **метод множників Лагранжа** та **метод поділу відрізка навпіл**, як приклади методів першої та другої групи.

2. Метод множників Лагранжа.

Метод множників Лагранжа є класичним методом знаходження екстремуму функції багатьох змінних у задачах нелінійного програмування з обмеженнями та без обмежень. Основні вимоги до задачі нелінійного програмування, яку можна розв'язувати методом множників Лагранжа:

- функція мети та обмеження описуються неперервно - диференційованими функціями;
- обмеження є рівняннями;
- відсутні вимоги невід'ємності та цілочисельності розв'язку.

Нехай задано задачу нелінійного програмування

$$F(x_1, x_2, ..., x_n) \to \max(\min) \tag{5.1}$$

$$g_i(x_1, x_2, .., x_n) = 0,\ i = \overline{1, m}, \tag{5.2}$$

де $F(x_1, x_2, ..., x_n)$, $g_i(x_1, x_2, .., x_n)$, $i = \overline{1, m}$, – неперервно-диференційовні функції в деякій області $D \in R^n$.

Для задачі (5.1) - (5.2) запишемо функцію

$$L(X, \lambda) = L(x_1, x_2, ..., x_n, \lambda_1, \lambda_2, ..., \lambda_n) = F(x_1, x_2, ..., x_n) + \sum_{i=1}^{m} \lambda_i g_i(x_1, x_2, ..., x_n),$$

яку називають **функцією Лагранжа**, де λ_1, λ_2, …, λ_m – множники Лагранжа.

Пошук розв'язку ведеться серед стаціонарних точок функції Лагранжа, тобто серед таких точок, в яких усі частинні похідні першого порядку функції $L(X, \lambda)$ дорівнюють нулю.

Знайдемо та прирівняємо до нуля частинні похідні $\dfrac{\partial L(X, \lambda)}{\partial x_j}$, $\dfrac{\partial L(X, \lambda)}{\partial \lambda_i}$, $j = \overline{1, n}$, $i = \overline{1, m}$. Одержимо систему рівнянь

$$\begin{cases} \dfrac{\partial F(x_1, x_2, ..., x_n)}{\partial x_j} + \sum\limits_{i=1}^{m} \lambda_i \dfrac{\partial g_i(x_1, x_2, ..., x_n)}{\partial x_j} = 0, \\ g_i(x_1, x_2, ..., x_n) = 0, \quad i = \overline{1, m}. \end{cases} \tag{5.3}$$

Розв'язуючи систему (5.3) отримуємо стаціонарні точки $X^0 = (x_1^0, x_2^0, ..., x_n^0)$, $\lambda^0 = (\lambda_1^0, \lambda_2^0, ..., \lambda_m^0)$ функції $L(X, \lambda)$, тобто точки підозрілі на екстремум. Питання про існування у знайдених точках екстремуму вирішується на основі достатніх умов екстремуму – дослідженні знаку другого диференціалу

$$d^2L(X,\lambda)=\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\partial^2 L}{\partial x_i \partial x_j}dx_i dx_j$$

в кожній стаціонарній точці $X^0=(x_1^0,x_2^0,...,x_n^0)$.

Достатні умови умовного екстремуму. Нехай $X^0=(x_1^0,x_2^0,...,x_n^0)$, $\lambda^0=(\lambda_1^0,\lambda_2^0,...,\lambda_m^0)$ є стаціонарною точкою функції Лагранжа $L(X,\lambda)$. Тоді

1. X^0 *є точкою умовного максимуму функції* $F(x_1,x_2,...,x_n)$ за умов (5.2), якщо $d^2L(X^0,\lambda^0)<0$ для довільних значень dx_1, …, dx_n, які всі одночасно не дорівнюють нулю і пов'язані між собою умовами

$$dg_i=\sum_{j=1}^{n}\frac{\partial g_i(X^0)}{\partial x_j}dx_j=0,\ i=\overline{1,m}. \qquad (5.4)$$

2. X^0 *є точкою умовного мінімуму функції* $F(x_1,x_2,...,x_n)$ за умов (5.2), якщо $d^2L(X^0,\lambda^0)>0$ для довільних значень dx_1, …, dx_n, які всі одночасно не дорівнюють нулю і пов'язані між собою умовами (5.4).
3. X^0 *не є точкою умовного екстремуму функції* $F(x_1,x_2,...,x_n)$ за умов (5.2), якщо $d^2L(X^0,\lambda^0)$ може набувати як додатних, так і від'ємних значень залежно від значень dx_1, …, dx_n, пов'язаних між собою умовами (5.4).
4. Питання про наявність умовного екстремуму в точці X^0 залишається відкритим, якщо не справджуються жодна з умов 1-3.

Зауваження. Система (5.3) в загальному випадку є системою з $n+m$ рівнянь та невідомих. Знаходження розв'язків такої системи є складною самостійною задачею.

Приклад 5.1. Дослідити функцію $F=F(x_1,x_2,x_3)=x_1x_2+x_2x_3$ на умовний екстремум методом множників Лагранжа при заданих обмеженнях: $x_1+x_2=2$, $x_2+x_3=2$.

Розв'язання.

Складемо функцію Лагранжа

$$L(X,\lambda)=L(x_1,x_2,x_3,\lambda_1,\lambda_2)=x_1x_2+x_2x_3+\lambda_1(x_1+x_2-2)+\lambda_2(x_2+x_3-2).$$

Знайдемо та прирівняємо до нуля частинні похідні $\dfrac{\partial L(X,\lambda)}{\partial x_j}$, $\dfrac{\partial L(X,\lambda)}{\partial \lambda_j}$.

Одержимо систему рівнянь:

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = x_2 + \lambda_1 = 0, \\ \dfrac{\partial L}{\partial x_2} = x_1 + x_3 + \lambda_1 + \lambda_2 = 0, \\ \dfrac{\partial L}{\partial x_3} = x_2 + \lambda_2 = 0, \\ \dfrac{\partial L}{\partial \lambda_1} = x_1 + x_2 - 2 = 0, \\ \dfrac{\partial L}{\partial \lambda_2} = x_2 + x_3 - 2 = 0. \end{cases}$$

Розв'язуємо систему, наприклад методом виключення, отримаємо стаціонарну точку $x_1 = x_2 = x_3 = 1$, при $\lambda_1 = \lambda_2 = -1$. Для того, щоб дослідити стаціонарну точку на екстремум, знайдемо другий диференціал функції Лагранжа:

$$\frac{\partial^2 L}{\partial x_1^2} = \frac{\partial^2 L}{\partial x_2^2} = \frac{\partial^2 L}{\partial x_3^2} = 0; \quad \frac{\partial^2 L}{\partial x_1 \partial x_2} = 1; \quad \frac{\partial^2 L}{\partial x_1 \partial x_3} = 0; \quad \frac{\partial^2 L}{\partial x_2 \partial x_3} = 1;$$

$$d^2 L = 2dx_1 dx_2 + 2dx_2 dx_3.$$

Умови (5.4) матимуть вигляд

$$dx_1 + dx_2 = 0;$$
$$dx_2 + dx_3 = 0;$$

звідки

$$dx_1 = -dx_2;$$
$$dx_3 = -dx_2.$$

Тоді

$$d^2 L = -2dx_2 dx_2 - 2dx_2 dx_2 = -4dx_2^2 < 0.$$

Отже, функція $F = F(x_1, x_2, x_3)$ досягає в точці $x_1 = x_2 = x_3 = 1$ свого умовного максимуму, $F_{\max} = 2$.

Зауваження. Якщо деякі змінні можна явно виразити через решту змінних з рівнянь (5.2) то **методом виключення** задачу можна спростити: звести до задачі безумовної оптимізації, зменшити кількість невідомих.

Наприклад, для попередньої задачі з рівнянь - обмежень маємо:

$$x_1 = 2 - x_2,$$
$$x_3 = 2 - x_2,$$

тому

$$F = (2 - x_2)x_2 + x_2(2 - x_2) = 4x_2 - 2x_2^2.$$

Отримали функцію однієї змінної $F = F(x_2)$, для знаходження екстремуму якої скористаємось методами диференціального числення. Знайдемо стаціонарні точки $F = F(x_2)$:

$$F' = 4 - 4x_2 = 0 \Rightarrow x_2 = 1.$$

Дослідимо знак похідної зліва й справа від стаціонарної точки (рис. 5.1):

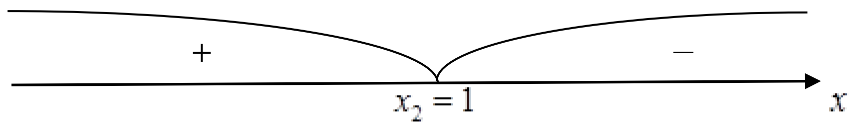

Рис. 5.1.

Отже, точка $x_2 = 1$ є точкою локального максимума функції $F = F(x_2)$. При $x_2 = 1$ отримаємо $x_1 = 1$, $x_3 = 1$. Таким чином, функція $F = F(x_1, x_2, x_3)$ досягає в точці $x_1 = x_2 = x_3 = 1$ свого умовного максимуму, $F_{\max} = 2$.

3. Метод поділу відрізка навпіл.

Методом поділу відрізка навпіл можна знайти наближений розв’язок задачі однопараметричної безумовної оптимізації

$$F(x) \to \max(\min),$$

якщо відомий відрізок невизначеності $[a;b]$, на якому функція однієї змінної $y = F(x)$ має тільки один максимум (мінімум).

Ідея методу поділу відрізка навпіл полягає у поступовому звуженні відрізка невизначеності до тих пір, поки із заданою точністю Δ не буде знайдена точка екстремуму.

Алгоритм методу поділу відрізка навпіл.

Нехай задано інтервал невизначеності $I_0 = [a;b]$ та необхідна точність Δ пошуку екстремуму. Позначимо $I_{i-1} = [x_L; x_R]$ інтервал невизначеності на *i*-му кроці (на нульовому кроці $x_L = a$ і $x_R = b$).

Крок *i*. Обчислюємо

$$x_1 = \frac{x_R + x_L - \Delta}{2}; \; x_2 = \frac{x_R + x_L + \Delta}{2}.$$

1. Якщо $F(x_1) > F(x_2)$, то точка максимуму x^* лежить між x_L та x_2. Тоді покладаємо $x_R = x_2$ і одержимо новий інтервал невизначеності $I_i = [x_L; x_2]$ (рис. 5.2).

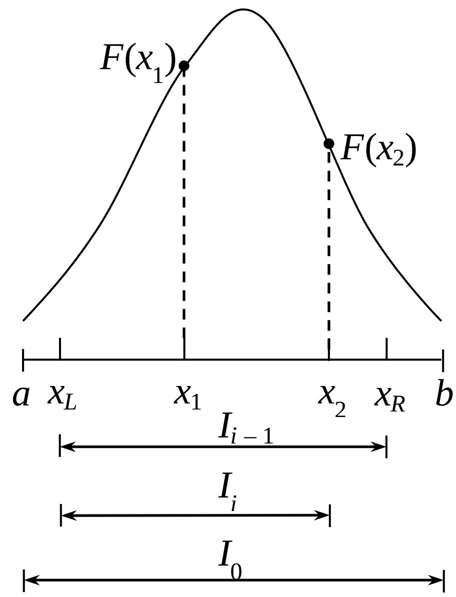

Рис. 5.2.

2. Якщо $F(x_1) < F(x_2)$, то точка максимуму x^* лежить між x_1 та x_R. Тоді покладаємо $x_L = x_1$ і $I_i = [x_1; x_R]$ (рис. 5.3).

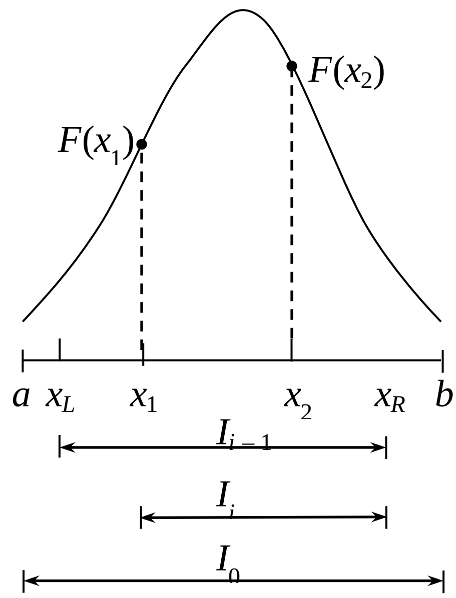

Рис. 5.3.

3. Якщо $F(x_1) = F(x_2)$, то точка максимуму x^* лежить між x_1 та x_2. Тоді покладаємо $x_L = x_1$, $x_R = x_2$ і $I_i = [x_1; x_2]$.

Оскільки на кожному кроці гарантовано $I_i < I_{i-1}$, в результаті звужується інтервал, в якому знаходиться точка максимуму функції $F(x)$.

Алгоритм закінчуємо на *k*-му кроці, коли довжина відрізка $I_k < \Delta$.

Зауваження. Алгоритм легко реалізувати програмними засобами, проте цей метод вимагає досить великої кількості обчислень. Значно швидше знаходить розв’язок метод золотого перерізу, який відрізняється лише формулою для обчислення $x_1 = x_R - \frac{\sqrt{5}-1}{2}(x_R - x_L)$ та $x_2 = x_L + \frac{\sqrt{5}-1}{2}(x_R - x_L)$.

4. Розв'язання задач нелінійного програмування в пакеті Maple.

Для розв'язання задач нелінійного програмування в ***Maple*** у пакеті **Optimization** призначені функції

NLPSolve(obj, constr, bd, opts)

яка має наступні параметри:
obj – алгебраїчна цільова функція;
constr – множина обмежень;
bd – послідовність у вигляді **name = range,** яка задає межі змінних;
opts – рівняння у вигляді **option = value,** де **option** може приймати значення **assume**, **feasibilitytolerance**, **iterationlimit**, **maximize**, **method**, **optimalitytolerance** або **output**;

QPSolve(obj, constr, bd, opts)

яка має наступні параметри:
obj – алгебраїчна, квадратична цільова функція;
constr – множина обмежень;
bd – послідовність у вигляді **name = range,** яка задає межі змінних;
opts – рівняння у вигляді **option = value,** де **option** може приймати значення **assume**, **feasibilitytolerance**, **iterationlimit**, **maximize** або **output**.

Задаючи параметри для функції ***NLPSolve*** можна вибрати метод, за яким буде шукатись розв'язок задачі нелінійного програмування. Більш детально з можливостями та параметрами функції пакета **Optimization** можна ознайомитись у довіднику системи ***Maple***.

За допомогою функції ***LagrangeMultipliers*** пакета **Student[MultivariateCalculus]** можна знайти стаціонарні точки методом множників Лагранжа.

Приклад 5.2. Методом множників Лагранжа знайти екстремум функції $F = 3x_1 x_2 + 4$, при умові $\frac{x_1^2}{8} + \frac{x_2^2}{2} = 100$.

Розв'язання.

Розв'яжемо задачу в ***Maple*** (рис. 5.4).

> restart :

Підключаємо пакет

> with(Student[MultivariateCalculus]) :

Задаємо функцію мети F та обмеження g

> $F := 3 \cdot x_1 \cdot x_2 + 4;\ g := \frac{x_1^2}{8} + \frac{x_2^2}{2} - 100;\ X := [x_1, x_2];$

$$F := 3\,x_1\,x_2 + 4$$

$$g := \frac{1}{8}\,x_1^2 + \frac{1}{2}\,x_2^2 - 100$$

$$X := [x_1, x_2] \qquad (1)$$

Знаходимо стаціонарні точки функції Лагранжа

> $LagrangeMultipliers(F, [g], X);$

$$[20, 10], [-20, -10], [-20, 10], [20, -10] \qquad (2)$$

> $LagrangeMultipliers(F, [g], X, output = detailed);$

$$[x_1 = 20, x_2 = 10, \lambda_1 = 6, 3\,x_1\,x_2 + 4 = 604], [x_1 = -20, x_2 = -10, \lambda_1 = 6, 3\,x_1\,x_2 + 4 = 604], [x_1 = -20, x_2 = 10, \lambda_1 = -6, 3\,x_1\,x_2 + 4 = -596], [x_1 = 20, x_2 = -10, \lambda_1 = -6, 3\,x_1\,x_2 + 4 = -596] \qquad (3)$$

Рис. 5.4.

Отже, маємо чотири стаціонарні точки. Зауважимо, що детальний розв'язок отримано за допомогою опції **output = detailed.**

Для функції двох змінних розв'язок можна відобразити графічно (рис. 5.5).

> $LagrangeMultipliers(F, [g], X, output = plot, showconstraints = false);$

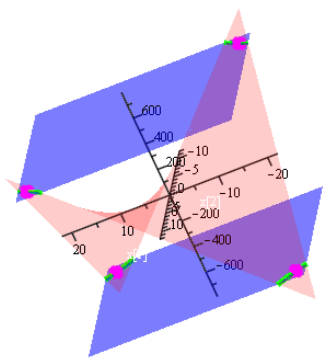

Рис. 5.5.

Отже, $F_{\max} = F(-20,-10) = F(20,10) = 604$,
$F_{\min} = F(-20,10) = F(20,-10) = -596$.

Приклад 5.3. Розвя'зати задачу нелінійного програмування

$$F = -x_1^2 - x_2^2 + 8x_1 + 10x_2 \to \max,$$

$$6 - 3x_1 - 2x_2 \ge 0,$$

$$x_1 \ge 0, \quad x_2 \ge 0.$$

Розв'язання.

Розв'яжемо задачу в ***Maple***. (рис. 5.6).

```
> restart :
```
Підключаємо пакет
```
> with(Optimization) :
```
Задаємо функцію мети F та обмеження g
```
> F := -x[1]^2 - x[2]^2 + 8·x[1] + 10·x[2]; g := 6 - 3·x[1] - 2·x[2];
```
$$F := -x_1^2 - x_2^2 + 8\,x_1 + 10\,x_2$$
$$g := 6 - 3\,x_1 - 2\,x_2 \qquad (1)$$

Рис. 5.6.

Знаходимо максимум функції мети (опція **maximize**) для невід’ємних значень змінних (опція **assume = nonnegative**) (рис. 5.7).

```
> NLPSolve(F, {g ≥ 0}, assume = nonnegative, maximize);
```
$$[21.3076923076923066,\ [x_1 = 0.307692307692307986,\ x_2 = 2.53846153846153788]] \qquad (2)$$

Рис. 5.7.

Отже, $F_{\max} = F(0{,}31; 2{,}54) \approx 21{,}31.$

Контрольні запитання.

1. Сформулюйте задачу нелінійного програмування.
2. Як класифікують методи нелінійного програмування?
3. Сформулюйте ідею методу поділу відрізка навпіл.
4. Які задачі можна розв’язувати методом поділу відрізка навпіл?
5. Які недоліки має метод поділу відрізка навпіл?
6. Сформулюйте основні вимоги до задачі нелінійного програмування, яку можна розв’язувати методом множників Лагранжа.
7. Опишіть метод множників Лагранжа.
8. Які недоліки має метод множників Лагранжа?

Завдання на самопідготовку.

1. Віртуальний університет ЛДУ БЖД [Електронний ресурс]. — Режим доступу: http://virt.ldubgd.edu.ua/
2. Системний аналіз та теорія прийняття рішень [Електронний ресурс] / Чмир Оксана Юріївна. — Режим доступу: http://virt.ldubgd.edu.ua/course/view.php?id=1765

Доцент
кафедри прикладної математики і механіки Оксана Чмир

Лекція обговорена на засіданні
кафедри прикладної математики і механіки

Протокол № від “ ” серпня 20 р.