Math 7502

Homework 5

Due: February 21, 2008

1. Consider the linear program (P)

$$\begin{array}{lccccccccc} \text{minimize} & 2x_1 & + & 5x_2 & + & 3x_3 & +5x_4 & +3x_5 & \\ \text{subject to} & x_1 & + & 3x_2 & + & x_3 & +2x_4 & +3x_5 & \geq 4 \\ & 2x_1 & + & 2x_2 & - & 2x_3 & +3x_4 & +x_5 & \geq 3 \\ & x_1 & , & x_2 & , & x_3, & x_4, & x_5 & \geq 0. \end{array}$$

∗ (i) Write down the dual program (D).

$$\begin{array}{lcccc} \text{maximize} & 4y_1 & + & 3y_2 & \\ \text{subject to} & y_1 & + & 2y_2 & \leq 2 \\ & 3y_1 & + & 2y_2 & \leq 5 \\ & y_1 & - & 2y_2 & \leq 3 \\ & 2y_1 & + & 3y_2 & \leq 5 \\ & 3y_1 & + & y_2 & \leq 3 \\ & y_1 & , & y_2 & \geq 0. \end{array}$$

In matrix form the primal program is

$$\begin{array}{ll} \text{minimize} & c^t \cdot x \\ \text{subject to} & Ax \geq b, x \geq 0, \end{array}$$

where $c^t = (2, 5, 3, 5, 3)$, $x^t = (x_1, x_2, x_3, x_4, x_5)$, $b^t = (4, 3)$ and

$$A = \begin{pmatrix} 1 & 3 & 1 & 2 & 3 \\ 2 & 2 & -2 & 4 & 1 \end{pmatrix}.$$

Using slack variables the program is

$$\begin{array}{lcccccccccc} \text{minimize} & 2x_1 & + & 5x_2 & + & 3x_3 & +5x_4 & +3x_5 & & & \\ \text{subject to} & x_1 & + & 3x_2 & + & x_3 & +2x_4 & +3x_5 & -z_1 & & = 4 \\ & 2x_1 & + & 2x_2 & - & 2x_3 & +3x_4 & +x_5 & & -z_2 & = 3 \\ & x_1 & , & x_2 & , & x_3, & x_4, & x_5, & z_1, & z_2 & \geq 0. \end{array}$$

Using slack variables the dual program is

$$\begin{array}{lccccccccc} \text{maximize} & 4y_1 & + & 3y_2 & & & & & & \\ \text{subject to} & y_1 & + & 2y_2 & +w_1 & & & & & = 2 \\ & 3y_1 & + & 2y_2 & & +w_2 & & & & = 5 \\ & y_1 & - & 2y_2 & & & +w_3 & & & = 3 \\ & 2y_1 & + & 3y_2 & & & & +w_4 & & = 5 \\ & 3y_1 & + & y_2 & & & & & +w_5 & = 3 \\ & y_1 & , & y_2, & w_1, & w_2, & w_3, & w_4, & w_5 & \geq 0. \end{array}$$

The dual program in matrix form is

$$\begin{array}{ll} \text{maximize} & b^t \cdot y \\ \text{subject to} & A^t y \leq c, y \geq 0. \end{array}$$

$*$ (ii) Solve the dual program (D) using the simplex method. We have for the dual program the simplex tableau

1	2	1	0	0	0	0	2
3	2	0	1	0	0	0	5
1	-2	0	0	1	0	0	3
2	3	0	0	0	1	0	5
3	1	0	0	0	0	1	3
4	3	0	0	0	0	0	0

This gives as basic feasible solution $(0, 0, 2, 5, 3, 5, 3)$. This is not optimal, due to the positive entries on the last row. We decide to include y_1 in the basic variables, as $4 > 3$ (Dantzig's rule). The choice of y_2 to be included is also possible. Since $3/3 < 5/3 < 2/1 < 5/2 < 3/1$ we choose to pivot on the $a_{51} = 3$ entry. We divide the fifth row by 3 and then subtract the resulting row from the first and third, twice the fifth row from the fourth and 3 times from the second. We also subtract 4 times the fifth row from the last row. This process gives successively the tableaux

1	2	1	0	0	0	0	2
3	2	0	1	0	0	0	5
1	-2	0	0	1	0	0	3
2	3	0	0	0	1	0	5
1	1/3	0	0	0	0	1/3	1
4	3	0	0	0	0	0	0

0	5/3	1	0	0	0	-1/3	1
0	1	0	1	0	0	-1	2
0	-7/3	0	0	1	0	-1/3	2
0	7/3	0	0	0	1	-2/3	3
1	1/3	0	0	0	0	1/3	1
0	5/3	0	0	0	0	-4/5	-4

Our basic solution is now $(1, 0, 1, 2, 2, 3, 0)$ and is not optimal, since we have a positive coefficient on the last row, $5/3$. We decide to include y_2 in the basic variables. Since

$$\frac{1}{5/3} < \frac{3}{7/3} < \frac{2}{1} < \frac{1}{1/3},$$

we choose to pivot on the $a_{12} = 5/3$ entry. We multiply the first row by 3/5. Then we subtract it from the second, multiply by 7/3 and add to the third and subtract from the fourth, mutiply it with 1/3 and subtract from the fifth, mutiply with 5/3 and subtract from the last row. We get the tableau

0	1	3/5	0	0	0	-1/5	3/5
0	1	0	1	0	0	-1	2
0	-7/3	0	0	1	0	-1/3	2
0	7/3	0	0	0	1	-2/3	3
1	1/3	0	0	0	0	1/3	1
0	5/3	0	0	0	0	-4/5	-4

0	1	3/5	0	0	0	-1/5	3/5
0	0	-3/5	1	0	0	-4/5	7/5
0	0	7/5	0	1	0	-4/5	17/5
0	0	-7/5	0	0	1	-1/5	8/5
1	0	-1/5	0	0	0	2/5	4/5
0	0	-1	0	0	0	-1	-5

The basic feasible solution is $(4/5, 3/5, 0, 7/5, 17/5, 8/5, 0)$ and is optimal as all the entries on the last row are negative or zero. The optimal value is $+5 = 4 \cdot 4/5 + 3 \cdot 3/5$.

$*$ (iii) Solve the dual program graphically.

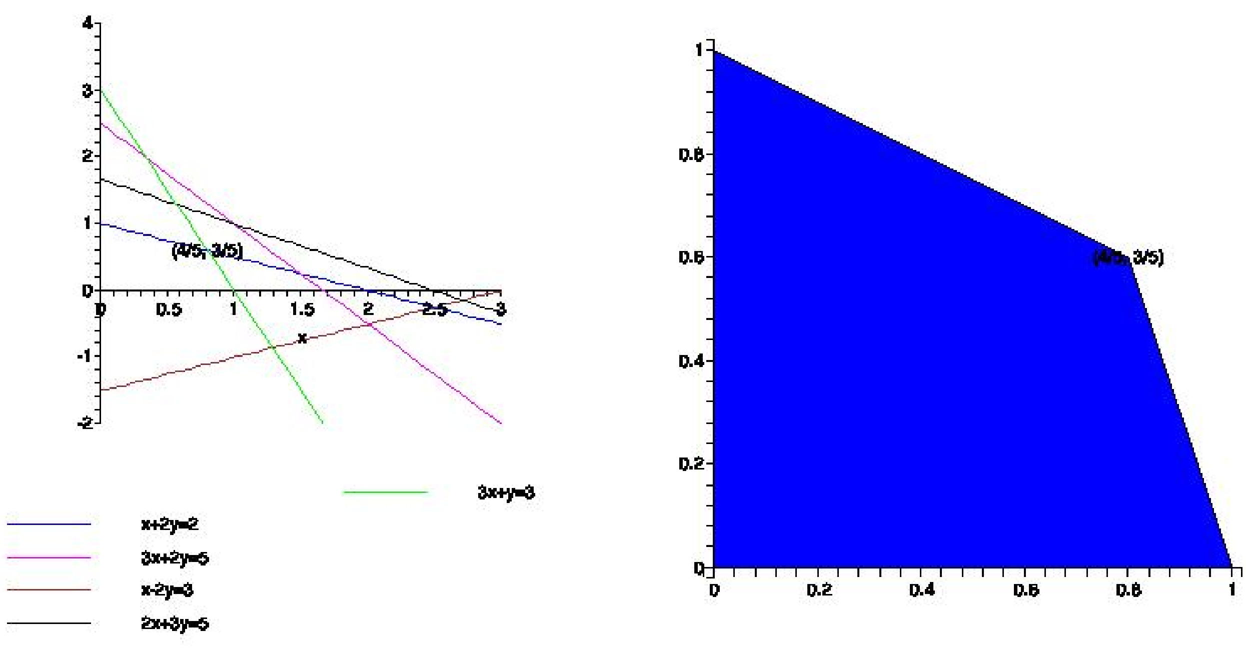

Figure 1: The feasible region for the dual program

$*$ (iv) Use complementary slackness to find an optimal solution to the primal program (P).

The solution to the dual program is $(y_1, y_2, w_1, w_2, w_3, w_4, w_5) = (4/5, 3/5, 0, 7/5, 17/5, 8/5, 0)$. Complementary slackness means that whenever $w_i \neq 0$ we must have $x_i = 0$. So $x_2 = x_3 = x_4 = 0$ at the optimal solution of the primal problem. Also $y_i \neq 0 \Longrightarrow z_i = 0$,

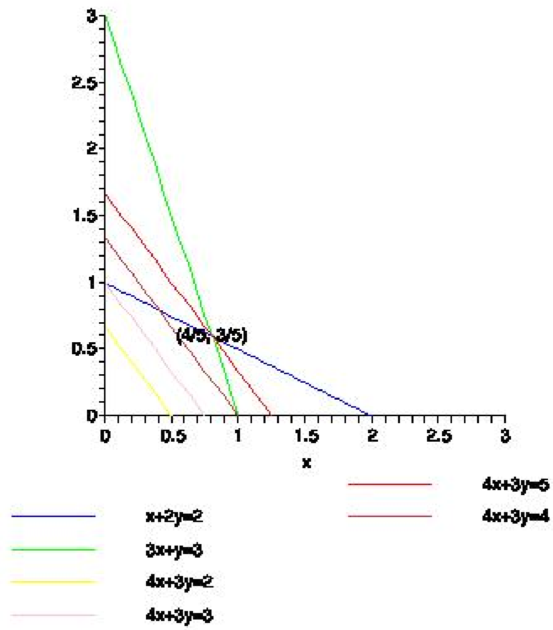

Figure 2: Graphical solution for the dual program

by complementary slackness. As a result $z_1 = z_2 = 0$. So the optimal for the primal program occurs where

$$x_1 + 3x_5 = 4, \quad 2x_1 + x_5 = 3.$$

This system has solution $x_1 = x_5 = 1$. The minimum of the objective function is

$$f = 2 \cdot 1 + 5 \cdot 0 + 3 \cdot 0 + 5 \cdot 0 + 3 \cdot 1 = 2 + 3 = 5,$$

and it agrees with the maximum for the dual program. This is the strong duality theorem.

(v) Use the two-phase simplex method to solve the program (P). You should appreciate how much faster complementary slackness is.

Because we see $-I_{2\times 2}$ and positive coefficients on the right, we introduce two artificial variables r_1, r_2 (one for each $-$ sign). We minimize first $r_1 + r_2$, i.e. we maximize $-r_1 - r_2$. This leads to the tableau

1	3	1	2	3	-1	0	1	0	4
2	2	-2	3	1	0	-1	0	1	3
0	0	0	0	0	0	0	-1	-1	0

This is not a valid simplex tableau, as we have nonzero entries below the identity matrix. We add the first two rows to the last to get the following valid simplex tableau

1	3	1	2	3	-1	0	1	0	4
2	2	-2	3	1	0	-1	0	1	3
3	5	-1	5	4	-1	-1	0	0	7

The basic feasible solution is $(0,0,0,0,0,0,0,4,3)$ and is not optimal for the maximum of $-r_1 - r_2$, as we have positive entries on the last row. We decide to include x_4 in the basic variables. Since $3/3 < 4/2$ we pivot on the $a_{24} = 3$ entry. We perform the row operations to get

1	3	1	2	3	-1	0	1	0	4
2/3	2/3	-2/3	1	1/3	0	-1/3	0	1/3	1
3	5	-1	5	4	-1	-1	0	0	7

-1/3	5/3	7/3	0	7/3	-1	2/3	1	-2/3	2
2/3	2/3	-2/3	1	1/3	0	-1/3	0	1/3	1
-1/3	5/3	7/3	0	7/3	-1	2/3	0	-5/3	2

The basic feasible solution is now $(0,0,0,1,0,0,0,2,0)$ and is not optimal for the maximum of $-r_1 - r_2$ as we have positive coefficients on the last row. We decide to include x_3 in the basic variables, and, since we have only one positive coefficient in the third column, we pivot on the $a_{13} = 7/3$ entry. We perform the row operations to get

-1/7	5/7	1	0	1	-3/7	2/7	3/7	-2/7	6/7
2/3	2/3	-2/3	1	1/3	0	-1/3	0	1/3	1
-1/3	5/3	7/3	0	7/3	-1	2/3	0	-5/3	2

-1/7	5/7	1	0	1	-3/7	2/7	3/7	-2/7	6/7
4/7	8/7	0	1	1	-2/7	-1/7	2/7	1/7	11/7
0	0	0	0	0	0	0	-1	-1	0

The basic feasible solution is $(0,0,6/7,11/7,0,0,0,0,0)$ and is optimal for the maximum of $-r_1 - r_2$. This maximum is 0. So the original program has nonempty feasible region and we erase the two columns of the artificial variables. We now minimize the original objective function $2x_1 + 5x_2 + 3x_3 + 5x_4 + 3x_5$. We get the tableau

-1/7	5/7	1	0	1	-3/7	2/7	6/7
4/7	8/7	0	1	1	-2/7	-1/7	11/7
-2	-5	-3	-5	-3	0	0	0

This is not a valid simplex tableau, since we have nonzero entries below the identity matrix. We add three times the first row and five times the second row to the third row to get the following valid simplex tableau

-1/7	5/7	1	0	1	-3/7	2/7	6/7
4/7	8/7	0	1	1	-2/7	-1/7	11/7
3/7	20/7	0	0	5	-19/7	1/7	73/7

The basic feasible solution is $(0, 0, 6/7, 11/7, 0, 0, 0)$ and is not optimal as we have positive coefficients on the last row. We decide to include x_2 in the basic variables, since $20/7 > 3/7$ (Dantzig's rule). Since $(6/7)/(5/7) < (11/7)/(8/7)$ we pivot on the $a_{12} = 5/7$ entry. We get the following tableau by performing the row operations

-1/5	1	7/5	0	7/5	-3/5	2/5	6/5
4/7	8/7	0	1	1	-2/7	-1/7	11/7
3/7	20/7	0	0	5	-19/7	1/7	73/7

-1/5	1	7/5	0	7/5	-3/5	2/5	6/5
4/5	0	-8/5	1	-3/5	2/5	-3/5	1/5
1	0	-4	0	1	-1	-1	7

The basic feasible solution is $(0, 6/5, 0, 1/5, 0, 0, 0)$ and is not optimal, since we have 1 in the last row. We pivot on the $a_{21} = 4/5$ entry. This gives the simplex tableau by performing the row operations

-1/5	1	7/5	0	7/5	-3/5	2/5	6/5
1	0	-2	5/4	-3/4	1/2	-3/4	1/4
1	0	-4	0	1	-1	-1	7

0	1	1	1/4	5/4	-1/2	1/4	5/4
1	0	-2	5/4	-3/4	1/2	-3/4	1/4
0	0	-2	-5/4	7/4	-3/2	-1/4	27/4

The basic feasible solution is $(1/4, 5/4, 0, 0, 0, 0, 0)$ and is not optimal. We include x_5 in the basic variables and pivot on the $a_{15} = 5/4$ entry. We perform the row operations to get

0	4/5	4/5	1/5	1	-2/5	1/5	1
1	0	-2	5/4	-3/4	1/2	-3/4	1/4
0	0	-2	-5/4	7/4	-3/2	-1/4	27/4

0	4/5	4/5	1/5	1	-2/5	1/5	1
1	3/5	-7/5	7/5	0	1/5	-3/5	1
0	-7/5	-17/5	-8/5	0	-4/5	-3/5	5

The basic feasible solution is $(1, 0, 0, 0, 1, 0, 0)$ and is optimal, since all the coefficients in the last row are nonpositive. The optimal value for the maximum of $-(2x_1+5x_2+3x_3+5x_4+3x_5)$ is -5, i.e. the optimal value for the minimum of $2x_1+5x_2+3x_3+5x_4+3x_5$ is 5.

$*$ (vi) Assume that the primal program corresponds to a diet problem: We have 5 types of food A_1, A_2, A_3, A_4, A_5 providing two types of nutrients C and M. The following table summarizes the nutritional content of the types of food, their cost, and the daily requirements for a healthy diet. We ignore units.

	A_1	A_2	A_3	A_4	A_5	Daily need
C	1	3	1	2	3	4
M	2	2	-2	3	1	3
Cost	2	5	3	5	3	

Notice that the negative number can be interpreted as follows: not only A_3 does not provide nutrient M but it removes from the body 2 units of nutrient M.

Give an economic interpretation of the dual program. In particular interpret the weak duality theorem, the strong duality theorem and the complementary slackness theorem.

There is an alternative way to get the necessary nutrients C, M: instead of eating the food types A_i, $i = 1, 2, 3, 4, 5$, we can buy from the health store nutritional supplements, say pills, containing the nutrients C and M. How much are we willing to pay for the pills? How much can the store charge for them? Let their price be y_1 and y_2 per unit, respectively. The store wants to maximize the profit of selling them. We will by 4 units of C and 3 of M. The profit of the store is $4y_1 + 3y_2$. But we will buy the pills, only if it is cheaper than buying the food that provide the nutrients. Suppose we do not buy a unit of food item A_1. Then we will not pay i.e. save 2, while we will not get 1 unit of C and 2 units of M, for which we have to pay y_1+2y_2. So we must have $y_1 + 2y_2 \leq 2$. Similarly we get the other dual constraints: $3y_1 + 2y_2 \leq 5$, $y_1 - 2y_2 \leq 3$, $2y_1 + 3y_2 \leq 5$ and $3y_1 + y_2 \leq 3$.

Weak duality theorem: For a set of feasible prices of the pills and a set of feasible amounts of food to eat, we will always have that the cost of pills (= profit of the store) is less than or equal to the cost of the food. This is because the diet problem has $Ax \geq b$ and the dual $A^t y \leq c$. This makes sense, because you will buy the pills as substitute for a certain food A_i, if it will cost you cheaper than A_i, while you buy enough pills to satisfy your daily need. The aggregate result is that, you may decide not to buy any food and buy only the pills, if you will pay less for the pills in total.

Strong duality theorem: The optimal of the two methods is the same. The store will set the prices to be as high as possible, but so that they are competitive vs buying the food. You will have no advantage then in buying the food or the nutrients, as they will cost the same.

Complementary slackness: If $z_i > 0$ then $y_i = 0$: z_i represent the excess of nutrient i in the diet, and y_i the cost of buying it. Interpretation: If in the optimal diet there is an excess of the nutrient i vs the daily need, you will be willing to pay 0 for it as a nutritional supplement.

If $w_j > 0$, then $x_j = 0$: x_j is the amount of food item j to include in the diet, while w_j represent the savings of not buying item j and buying the nutritional supplements instead. Interpretation: If in the optimal diet there are savings in buying the supplements instead of item j, then you should do so, and buy 0 of food item j.

2. Recall the definition of a convex function: Given $f : \mathbf{R} \to \mathbf{R}$ and $x, y \in \mathbf{R}$, $t \in [0,1]$, we have

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

(i) Prove Jensen's inequality: Given nonnegative scalars λ_i, $i = 1, 2, \ldots, k$ with $\sum_{i=1}^{k} \lambda_i = 1$, and points x_i, $i = 1, \ldots, k$, we have for a convex function $f(x)$ the inequality

$$f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leq \sum_{i=1}^{k} \lambda_i f(x_i).$$

Hint: Use induction with a clever choice in the inductive step to produce $k-1$ nonnegative numbers with sum equal to 1.

We start the induction with $k = 2$. Since $\lambda_1 + \lambda_2 = 1$, we can set $\lambda_2 = t$ and $\lambda_1 = 1 - t$ in the definition of convexity to get

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2),$$

which is what we try to prove.

Now we assume that we proved the result for $k-1$ scalars λ_i, and points x_i, $i = 1, \ldots, k-1$. We try to prove it for k. Given k scalars λ_i, which are nonnegative and $\sum_{i=1}^{k} \lambda_i = 1$ and given k points x_i, $i = 1, \ldots, k$:

(i) If $\lambda_k = 1$, then all the other $\lambda_i = 0$, $i = 1, \ldots, k-1$, as they are nonnegative with sum 0. Jensen's inequality now has one term on each side, namely $f(x_k)$. So it is obvious.

(ii). If $\lambda_k \neq 1$, we introduce the numbers

$$\mu_i = \frac{\lambda_i}{1 - \lambda_k}, \quad i = 1, \ldots, k-1.$$

Since $0 \leq \lambda_k < 1$, we have that $\mu_i \geq 0$. On the other hand

$$\sum_{i=1}^{k-1} \mu_i = \frac{1}{1-\lambda_k} \sum_{i=1}^{k-1} \lambda_i = \frac{1-\lambda_k}{1-\lambda_k} = 1,$$

since $\displaystyle\sum_{i=1}^{k} \lambda_i = 1$. We apply Jensen's inequality for $k-1$, which is the inductive hypothesis to get

$$f\left(\sum_{i=1}^{k-1} \mu_i x_i\right) \leq \sum_{i=1}^{k-1} \mu_i f(x_i).$$

We substitute the values of μ_i on the right hand side and multiply with $1-\lambda_k$ to get

$$(1-\lambda_k) f\left(\sum_{i=1}^{k-1} \mu_i x_i\right) \leq \sum_{i=1}^{k-1} \lambda_i f(x_i).$$

We add to both sides the missing term on the right, i.e. $\lambda_k f(x_k)$ to get

$$\lambda_k f(x_k) + (1-\lambda_k) f\left(\sum_{i=1}^{k-1} \mu_i x_i\right) \leq \sum_{i=1}^{k} \lambda_i f(x_i).$$

We apply the definition of convexity to the two terms on the left to get

$$f\left(\lambda_k x_k + (1-\lambda_k)\sum_{i=1}^{k-1} \mu_i x_i\right) \leq \lambda_k f(x_k) + (1-\lambda_k) f\left(\sum_{i=1}^{k-1} \mu_i x_i\right) \leq \sum_{i=1}^{k} \lambda_i f(x_i).$$

The first term can now be simplified, by recalling the definition of $\mu_i = \lambda_i/(1-\lambda_k)$:

$$f\left(\lambda_k x_k + \sum_{i=1}^{k-1} \lambda_i x_i\right) = f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \leq \sum_{i=1}^{k} \lambda_i f(x_i).$$

This is simply Jensen's inequality with k terms.

(ii) Prove that if f is differentiable and $f'(x)$ is an increasing function, then f is convex.

Hint: Use the Mean Value Theorem.

Let $x, y \in \mathbf{R}$ with $x < y$ and z a point between them, i.e, we can find a $t \in (0,1)$ with

$$z = (1-t)x + ty, \quad x < z < y.$$

We apply the mean value theorem in the intervals $[x,z]$ and $[z,y]$ to prove that we can find two points $c \in (x,z)$, $d \in (z,y)$ such that

$$\frac{f(x)-f(z)}{x-z} = f'(c), \quad \frac{f(z)-f(y)}{z-y} = f'(d).$$

Since f' is increasing and $c < z < d$, we have $f'(c) \leq f'(d)$. This implies

$$\frac{f(x)-f(z)}{x-z} \leq \frac{f(z)-f(y)}{z-y} \Longrightarrow \frac{f(x)-f(z)}{t(x-y)} \leq \frac{f(z)-f(y)}{(1-t)(x-y)}$$

Since $x < y$ we conclude

$$(f(x) - f(z))(1-t) \geq t(f(z) - f(y)) \implies (1-t)f(x) - (1-t)f(z) \geq tf(z) - tf(y)$$

$$\implies (1-t)f(x) + tf(y) \geq (t+1-t)f(z) = f(z) = f((1-t)x + ty).$$

If z is one of the endpoints of the interval $[x, y]$, the result is obvious: e.g., if $z = x$, then $t = 0$, $f(z) = 1 \cdot f(x) + (1-1)f(y)$.

If $x = y$, the same applies.

(iii) Prove that $-\ln(x)$ is a convex function for $x > 0$.

With $f(x) = -\ln(x)$, we have $f'(x) = -1/x$. Since for $x > 0$, $1/x$ is decreasing, $f'(x)$ is increasing and we apply (ii) to get the result.