Course Materials for GEN-AI

Northeastern University

These materials have been prepared and sourced for the course **GEN-AI** at Northeastern University. Every effort has been made to provide proper citations and credit for all referenced works.

If you believe any material has been inadequately cited or requires correction, please contact me at:

Instructor: Ramin Mohammadi
r.mohammadi@northeastern.edu

Thank you for your understanding and collaboration.

Convex Conjugate Functions

1 Definition of Convex Conjugate

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper, convex, and lower semi-continuous function. The **convex conjugate** (also known as the **Fenchel conjugate**) of f is defined as:

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(\langle y, x \rangle - f(x) \right), \tag{1}$$

where $\langle y, x \rangle$ denotes the inner product (dot product) of y and x. Intuitively, $f^*(y)$ captures the maximum difference between the linear functional $\langle y, x \rangle$ and the function $f(x)$ for all x.

1.1 Geometric Intuition

The convex conjugate $f^*(y)$ can be seen as the "envelope" of all the tangent hyperplanes (or lines in the 2D case) to the graph of $f(x)$. It represents the steepest linear function that lies below the graph of $f(x)$ at a given slope y.

1.2 Key Properties of the Convex Conjugate

- **Duality:** $(f^*)^* = f$ if f is convex, proper, and lower semi-continuous.
- **Fenchel Inequality:** For all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$, we have

$$f(x) + f^*(y) \geq \langle x, y \rangle. \tag{2}$$

- **Convexity:** $f^*(y)$ is always a convex function, even if $f(x)$ is not strictly convex.

2 Examples

2.1 Example 1: Convex Conjugate of a Quadratic Function

Step 1: Definition of Function
Let $f(x) = \frac{1}{2}x^2$. We compute the conjugate using the definition:

$$f^*(y) = \sup_{x \in \mathbb{R}} \left(yx - \frac{1}{2}x^2 \right). \tag{3}$$

Step 2: Solve for x

To maximize $yx - \frac{1}{2}x^2$ with respect to x, we take the derivative and set it to zero:

$$\frac{d}{dx}\left(yx - \frac{1}{2}x^2\right) = y - x = 0 \implies x = y. \tag{4}$$

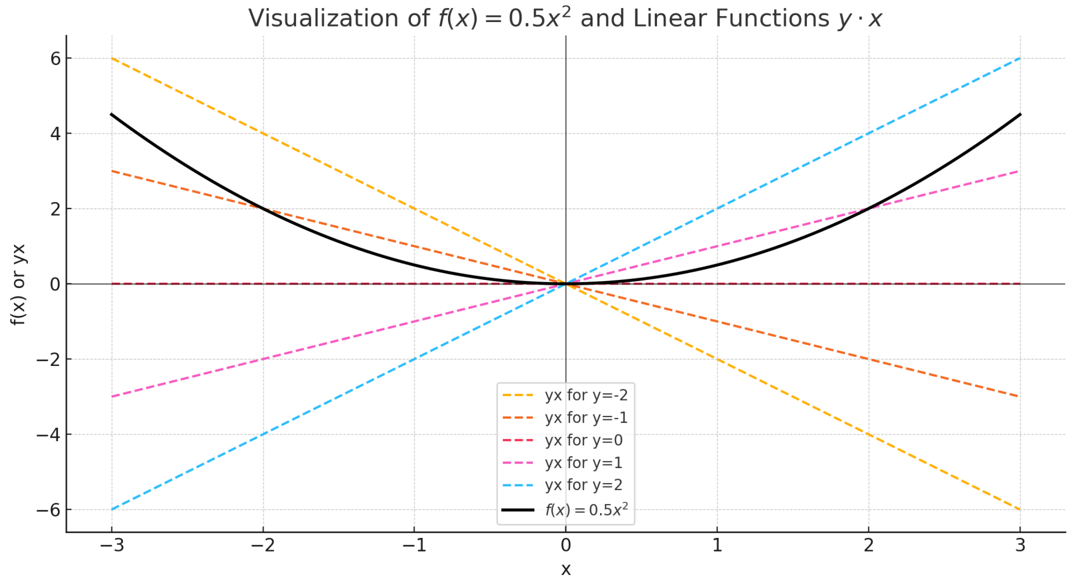


Figure 1: Visualization of $f(x) = \frac{1}{2}x^2$ and its conjugate $f^*(y) = \frac{1}{2}y^2$.

The black curve represents the function $f(x) = \frac{1}{2}x^2$. The dashed lines are the linear functions $y \cdot x$ for different values of y (like $y = -2, -1, 0, 1, 2$).

Step 3: Compute the Supremum Substituting $x = y$ into the objective function, we get:

$$f^*(y) = y \cdot y - \frac{1}{2}y^2 = \frac{1}{2}y^2. \tag{5}$$

Result: The conjugate of $f(x) = \frac{1}{2}x^2$ is:

$$f^*(y) = \frac{1}{2}y^2. \tag{6}$$

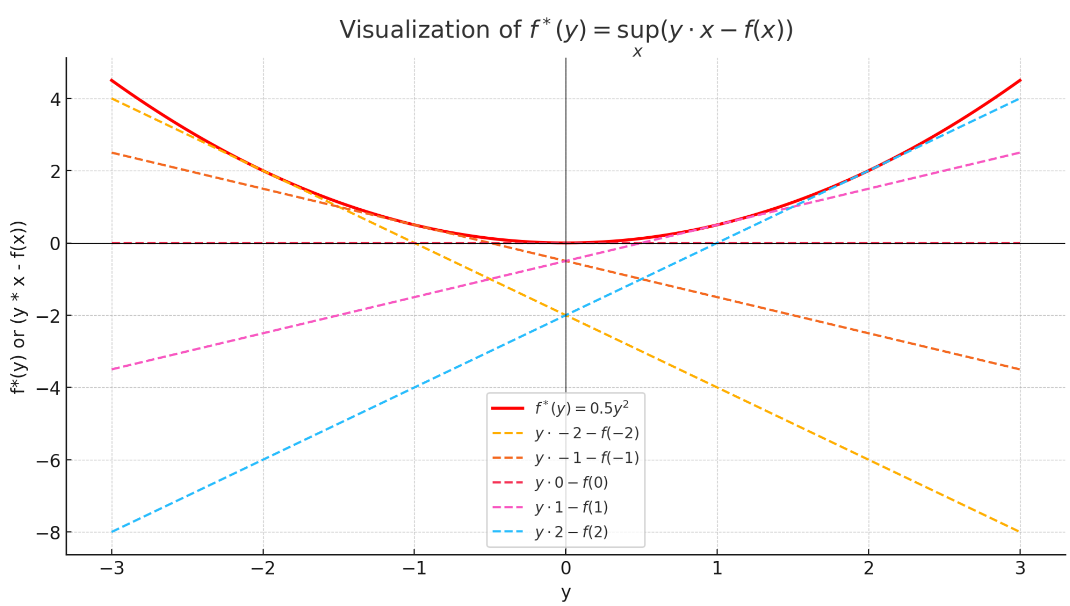


Figure 2: Visualization of $f^*(y)$

The red curve shows the conjugate $f^*(y) = \frac{1}{2}y^2$, which is the result of optimizing the supremum:

$$f^*(y) = \sup_x (y \cdot x - f(x)) \tag{7}$$

for every y.

The dashed curves show $y \cdot x - f(x)$ for different values of x (like $x = -2, -1, 0, 1, 2$). The curve $f^*(y)$ is the "upper envelope" of all these dashed lines, which means it's the supremum of these values for each y.

Fenchel Inequality

The heatmap visualizes the Fenchel inequality:

$$f(x) + f^*(y) \geq \langle x, y \rangle \implies f(x) + f^*(y) - x \cdot y \geq 0 \tag{8}$$

The quantity $f(x) + f^*(y) - x \cdot y$ (called the Fenchel gap) is plotted as a function of x and y.

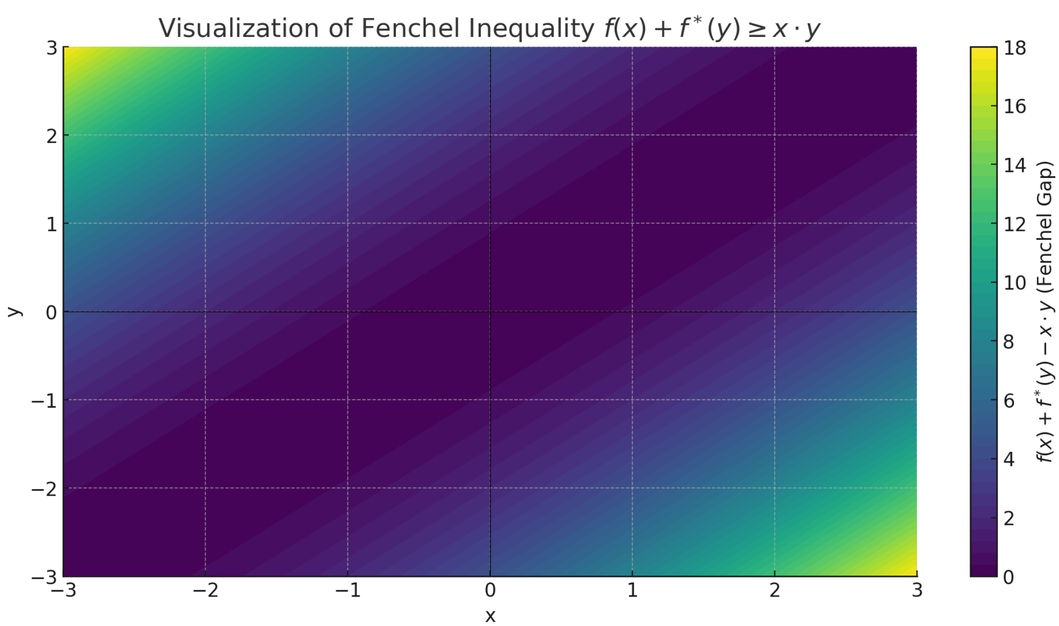
The Fenchel gap is:

$$\text{Fenchel gap} = f(x) + f^*(y) - x \cdot y \tag{9}$$

Using the specific forms of $f(x)$ and $f^*(y)$, we have:

$$\text{Fenchel gap} = \frac{1}{2}x^2 + \frac{1}{2}y^2 - x \cdot y \tag{10}$$

This is a quadratic function in x and y. We can rewrite it to understand its structure.



The heatmap shows the "Fenchel gap" $f(x) + f^*(y) - x \cdot y$ as a function of x and y.

By definition, Fenchel's inequality states that this quantity is always non-negative:

$$f(x) + f^*(y) \geq \langle x, y \rangle \tag{11}$$

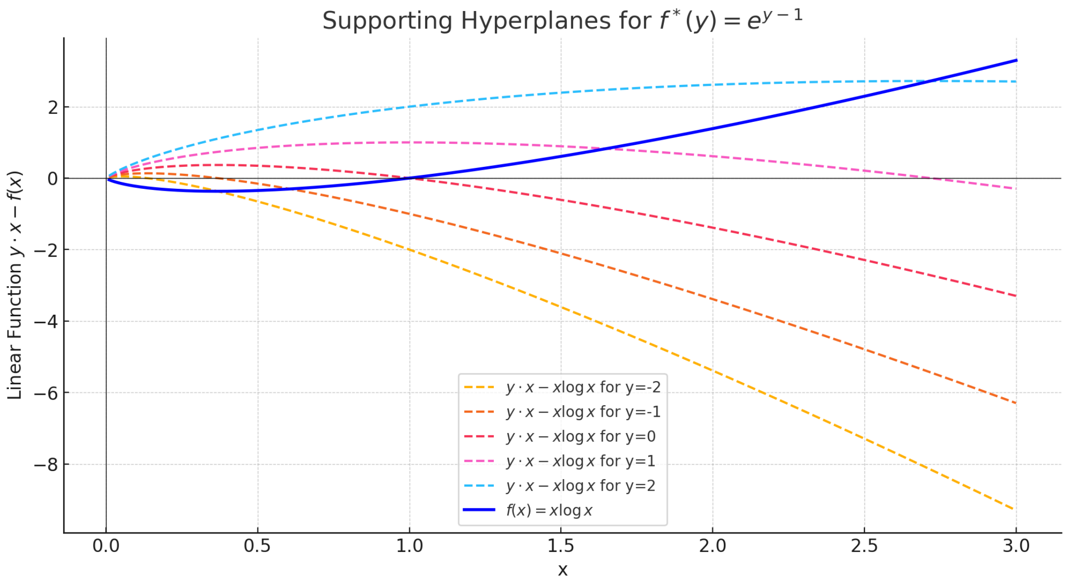
Darker regions correspond to smaller values, and lighter regions indicate larger values of the gap. This visualization illustrates that the Fenchel gap is always non-negative, in accordance with the Fenchel inequality.

Example 2: Convex Conjugate of Negative Entropy

Step 1: Definition of Function

Let $f(x) = x \log x$ (for $x > 0$ and $f(x) = +\infty$ for $x \leq 0$). We compute its conjugate:

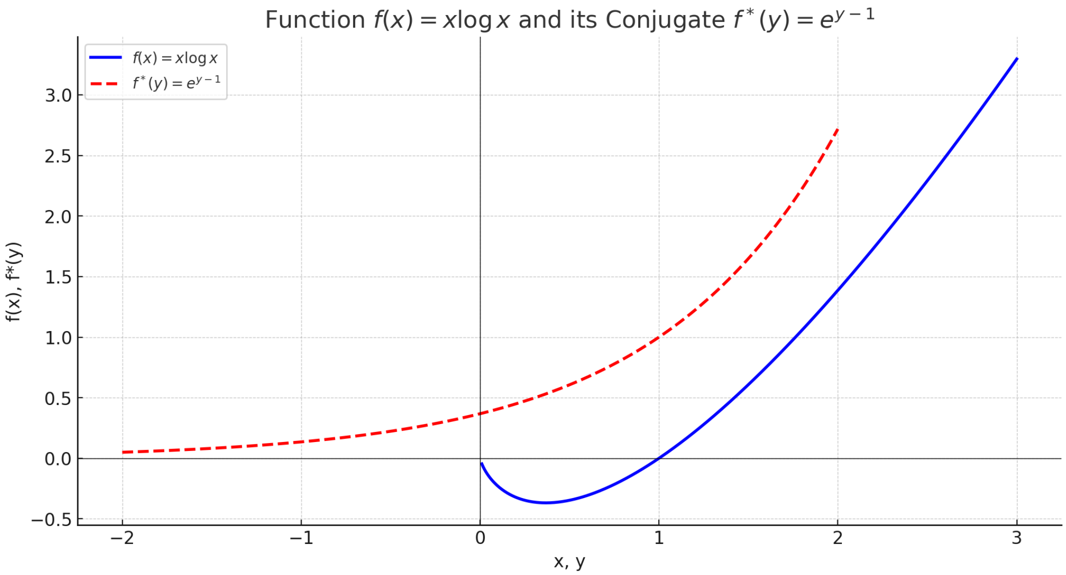
$$f^*(y) = \sup_{x>0} (yx - x \log x) . \tag{12}$$



Step 2: Solve for x

To maximize $yx - x \log x$, we compute the derivative and set it to zero:

$$\frac{d}{dx} (yx - x \log x) = y - \log x - 1 = 0 \implies x = e^{y-1}. \tag{13}$$



Step 3: Compute the Supremum

Substituting $x = e^{y-1}$ into the objective function:

$$f^*(y) = y \cdot e^{y-1} - e^{y-1} \log e^{y-1} = e^{y-1}(y - (y-1)) = e^{y-1}. \tag{14}$$

Result: The conjugate of $f(x) = x \log x$ is:

$$f^*(y) = e^{y-1}. \tag{15}$$