

Goog

Legendre and Legendre-Fenchel transforms

What do you need to know to understand this topic?

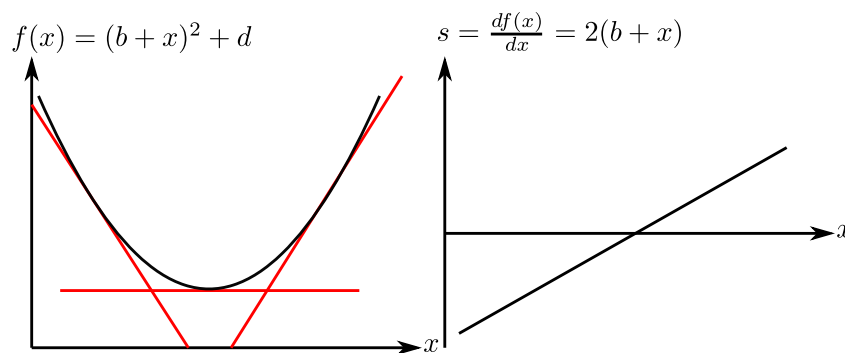
- Derivatives of a function
- What are convex and non-convex functions

Sections

- What is the Legendre Transform?
- What is the Legendre-Fenchel transform?
 - Examples

What is the Legendre Transform?

The Legendre transform is a transformation from a convex differentiable function $f(\mathbf{x})$ to a function that depends on the family of tangents $s = \nabla_{\mathbf{x}} f(\mathbf{x})$. To see how it works, you must first realize that for a convex and smooth function there is a one-to-one correspondence between \mathbf{x} and $s(\mathbf{x})$, because $s(\mathbf{x})$ is a monotonic function of \mathbf{x} , i.e., the derivative is always increasing with \mathbf{x} . You can see this notion in the plot below.



slope:

$$s = \left. \frac{df(\mathbf{x})}{d\mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}^*}$$

If we want a function of \mathbf{s} , we could first find the inverse of $\mathbf{s}(\mathbf{x})$, say $\mathbf{x}(\mathbf{s})$, and apply it to $f(\mathbf{x})$. We would then get $f(\mathbf{x}(\mathbf{s}))$. Of course there are many ways to get a function of \mathbf{s} , not necessarily the one presented right now. Instead, the Legendre transform is the expression:

$$f^*(\mathbf{s}) = \mathbf{s}^T \mathbf{x}(\mathbf{s}) - f(\mathbf{x}(\mathbf{s}))$$

To find out how this expression is derived, let's write the full expression of a tangent (see the picture below):

$$y = \mathbf{s}^T \mathbf{x} + b,$$

where b represents the value of the tangent when $\mathbf{x} = \mathbf{0}$. We know that at \mathbf{x}^* the tangent touches $f(\mathbf{x})$, therefore:

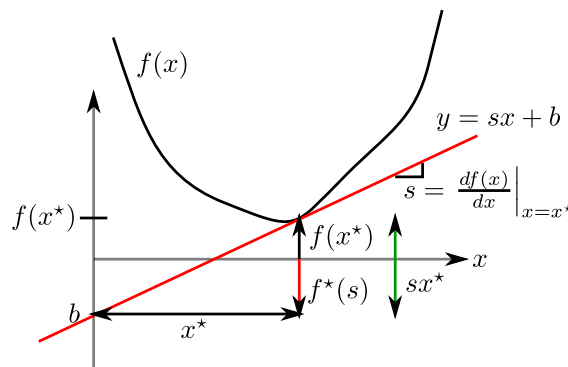
$$f(\mathbf{x}^*) = \mathbf{s}^T \mathbf{x}^* + b$$

$$-b = \mathbf{s}^T \mathbf{x}^* - f(\mathbf{x}^*)$$

Let's call the term $-b$, $f^*(\mathbf{s})$. The reason for it is that this term changes with \mathbf{x}^* and therefore with \mathbf{s} . Then, for any point \mathbf{x} :

$$f^*(\mathbf{s}) = \mathbf{s}^T \mathbf{x} - f(\mathbf{x})$$

In words, for any point $\mathbf{x} = \mathbf{x}^*$, the transform is the negative of parameter b of the tangent line at that point (the value of the tangent line when $\mathbf{x} = \mathbf{0}$). Graphically, this is depicted in figure below:



In a geometric sense, there is a triangle shown. In it, the slope of the tangent (s) times the horizontal side (x^*) equals the vertical side, which is the sum of $f(x)$ and $f^*(s)$.

Although this seems all a bit random, this transform has some nice properties. For one, we can apply it again to the transform to get back the original function:

$$\begin{aligned} f^{**}(\mathbf{x}) &= \mathbf{x}^T \mathbf{s} - f^*(\mathbf{s}) \\ &= \mathbf{x}^T \mathbf{s} - (\mathbf{s}^T \mathbf{x} - f(\mathbf{x})) \\ &= f(\mathbf{x}) \end{aligned}$$

In the same way that the slope of $f(\mathbf{x})$ is \mathbf{s} , we have that the slope of $f^*(\mathbf{s})$ is \mathbf{x} :

$$\frac{df^*(\mathbf{s})}{d\mathbf{s}} = \frac{d\mathbf{s}^T \mathbf{x} - f(\mathbf{x})}{d\mathbf{s}} = \mathbf{x}$$

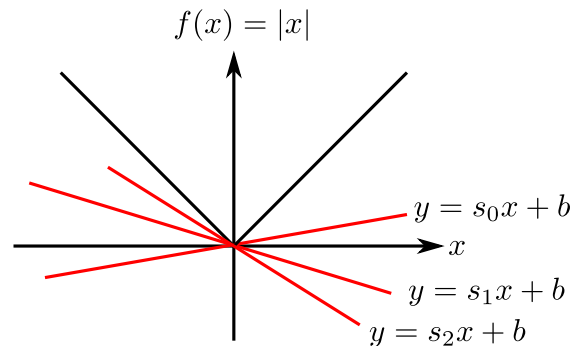
Say that we want to find the minimum of $f(\mathbf{x})$ and we have $f^*(\mathbf{s})$. Suffice to say that at the minimum, the slope of the tangent is zero. Therefore:

$$\begin{aligned} f^*(\mathbf{0}) &= \mathbf{0}^T \mathbf{x}_{min} - f(\mathbf{x}_{min}) \\ f(\mathbf{x}_{min}) &= -f^*(\mathbf{0}) \end{aligned}$$

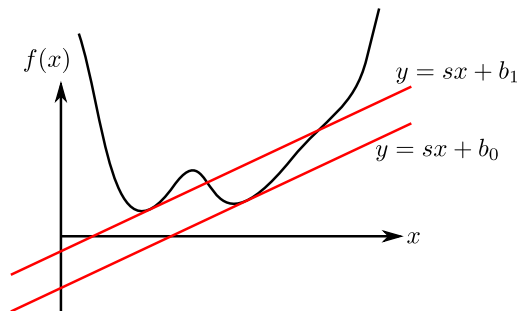
Have any suggestions for a new topic? Maybe I can wLet me kno

What is the Legendre-Fenchel transform?

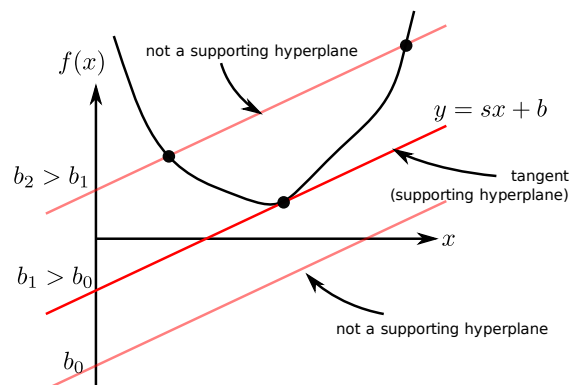
The Legendre transform is pretty useful on its own, but it is limited to convex and differentiable functions. If any of these properties fail, the transform cannot be used. Consider the case $f(x) = |x|$. The slope at $x = 0$ is not defined. All we know is that, at point $x = 0$, the subdifferential is between -1 and 1 ($s \in [-1, 1]$).



Also, non-convex functions have the same derivative for different values of \mathbf{x} , which makes it impossible to form a unique correspondence between \mathbf{x} and \mathbf{s} .



Here comes a nifty trick to generalize the Legendre transform to this type of functions. As you now know, the value $f^*(\mathbf{s})$ comes from finding b of the tangent, which is the smallest b term for a line (hyperplane for several dimensions) with slope \mathbf{s} that crosses the convex function $f(\mathbf{x})$. As shown in the picture below, the smallest b that crosses the $f(\mathbf{x})$ is the one from the tangent. Smaller than that and the line will not even touch $f(\mathbf{x})$. So, in a more general sense, we are finding all values of b with which lines with slope \mathbf{s} cross $f(\mathbf{x})$ and then find the minimum b .



This more general rule applies to non-differentiable or non-convex functions. So, when a line with slope \mathbf{s} crosses $f(\mathbf{x})$, we have:

$$\begin{aligned} \mathbf{s}^T \mathbf{x} + b &= f(\mathbf{x}) \\ b &= f(\mathbf{x}) - \mathbf{s}^T \mathbf{x} \end{aligned}$$

Have any suggestions for a new topic? Maybe I can write about it. Let me know.

$$\begin{aligned}
 b &= \inf_{\mathbf{x}} (f(\mathbf{x}) - \mathbf{s}^T \mathbf{x}) \\
 &= \inf_{\mathbf{x}} (-(\mathbf{s}^T \mathbf{x} - f(\mathbf{x}))) \\
 &= -\sup_{\mathbf{x}} (\mathbf{s}^T \mathbf{x} - f(\mathbf{x}))
 \end{aligned}$$

where we used the property $\inf(\mathbf{x}) = -\sup(-\mathbf{x})$. Since $f^*(\mathbf{s}) = -b$, we finally have:

$$f^*(\mathbf{s}) = -b = \sup_{\mathbf{x}} (\mathbf{s}^T \mathbf{x} - f(\mathbf{x}))$$

This is the Legendre-Fenchel transform, also known as convex conjugate. Note that now **the transform is not reversible**, i.e., you cannot get the original function by applying the transform to the transform. On the other hand, **the transform of the transform is convex, even if the original function is not.**

Examples

Linear function

For the linear function:

$$f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} - b$$

the Legendre transform is

$$\begin{aligned}
 &\sup_{\mathbf{x}} \mathbf{s}^T \mathbf{x} - f(\mathbf{x}) \\
 &\sup (\mathbf{s}^T - \mathbf{c}^T) \mathbf{x} + b
 \end{aligned}$$

For $\mathbf{s}^T \neq \mathbf{c}^T$, the supremum is found for infinite values of \mathbf{x} . Therefore, the condition $\mathbf{s}^T = \mathbf{c}^T$ is necessary to have a bounded conjugate:

$$f^*(\mathbf{s}) = \begin{cases} b & \mathbf{s}^T = \mathbf{c}^T \\ \infty & \text{otherwise} \end{cases}$$

Weighted Least squares

The weighted least squares problem is defined as:

$$f(\mathbf{x}) = \frac{1}{2} \|\mathbf{A}(\mathbf{y} - \mathbf{x})\|_2^2$$

The Legendre transform is then:

$$\sup_{\mathbf{x}} \mathbf{s}^T \mathbf{x} - \frac{1}{2} \|\mathbf{A}(\mathbf{y} - \mathbf{x})\|_2^2$$

To find the value of \mathbf{x} that maximizes the above expression, we equate to zero its derivative w.r.t. \mathbf{x} :

$$\mathbf{s} + \mathbf{A}^T \mathbf{A}(\mathbf{y} - \mathbf{x}) = \mathbf{0}$$

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{s} + \mathbf{y}$$

Now we replace \mathbf{x} in the original transform:

$$\begin{aligned}
 &\sup_{\mathbf{x}} \mathbf{s}^T \mathbf{x} - \frac{1}{2} \|\mathbf{A}(\mathbf{y} - \mathbf{x})\|_2^2 \\
 &\mathbf{s}^T \left((\mathbf{A}^T \mathbf{A})^{-1} \mathbf{s} + \mathbf{y} \right) - \frac{1}{2} \left\| \mathbf{A} \left(\mathbf{y} - (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{s} - \mathbf{y} \right) \right\|_2^2 \\
 &\mathbf{s}^T \left((\mathbf{A}^T \mathbf{A})^{-1} \mathbf{s} + \mathbf{y} \right) - \frac{1}{2} \left\| \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{s} \right\|_2^2
 \end{aligned}$$

Have any suggestions for a new topic? Maybe I can write about it. [Let me know](#)

The common case when \mathbf{A} is a diagonal matrix $\mathbf{\Lambda}$, the above expression can be simplified:

$$\frac{1}{2} \mathbf{s}^T (\mathbf{\Lambda})^{-2} \mathbf{s} + \mathbf{s}^T \mathbf{y} = \frac{1}{2} \|\mathbf{\Lambda}^{-1} \mathbf{s}\| + \mathbf{s}^T \mathbf{y}$$

When \mathbf{A} is the identity matrix, it becomes:

$$\frac{1}{2} \|\mathbf{s}\|_2^2 + \mathbf{s}^T \mathbf{y}$$

ℓ_1 -norm

For the ℓ_1 -norm:

$$f(\mathbf{x}) = a \|\mathbf{x}\|_1$$

the Legendre transform is

$$\sup \mathbf{s}^T \mathbf{x} - a \|\mathbf{x}\|_1$$

The transform can be written as the sum of elements of \mathbf{s} and \mathbf{x} as:

$$\sup \sum (s_i x_i - a |x_i|) = \sup \sum \left(\begin{cases} (s_i - a) x_i & , x_i \geq 0 \\ (s_i + a) x_i & , x_i < 0 \end{cases} \right)$$

This way is easy to see that if any element $s_i > a$, then any vector \mathbf{x} with $x_i \rightarrow \infty$ will push the transform to infinity. Likewise, if any element $s_i < -a$, then any vector \mathbf{x} with $x_i \rightarrow -\infty$ will push the transform to ∞ too. For any $-a \leq s_i \leq a$, $(s_i - a) \leq 0$ and $(s_i + a) \geq 0$. Then, under this condition, the expressions above are always lower than zero because they always multiply terms with opposite signs. That in turn, means that the supremum is found for $x_i = 0$. In conclusion, the transform is:

$$f^*(\mathbf{s}) = \sup_{\mathbf{x}} (\mathbf{s}^T \mathbf{x} - a \|\mathbf{x}\|_1) = \|\mathbf{s}\|_{\infty}^a = \begin{cases} 0 & |s_i| \leq a, \forall i \\ +\infty & \text{otherwise} \end{cases}$$

Weighted ℓ_1 -norm

For the weighted ℓ_1 -norm case:

$$f(\mathbf{x}) = a \|\mathbf{W}^{-1} \mathbf{x}\|_1$$

the Legendre transform is

$$\begin{aligned} \sup \mathbf{s}^T \mathbf{x} - a \|\mathbf{W}^{-1} \mathbf{x}\|_1 \\ \sup (\mathbf{s}^T \mathbf{W}) \mathbf{z} - a \|\mathbf{z}\|_1 \end{aligned}$$

with $\mathbf{z} = \mathbf{W}^{-1} \mathbf{x}$ and assuming matrix \mathbf{W} can be inverted. Following a similar approach to the previous example, the values for s_i are limited, but now each is scaled differently:

$$\sup \sum (\mathbf{s}^T \mathbf{w}_i z_i - a |z_i|) = \sup \sum \left(\begin{cases} (\mathbf{s}^T \mathbf{w}_i - a) z_i & , z_i \geq 0 \\ (\mathbf{s}^T \mathbf{w}_i + a) z_i & , z_i < 0 \end{cases} \right)$$

where \mathbf{w}_i is the i^{th} column of \mathbf{W} . For the supremum to be bounded, we must have that each $-a < \mathbf{s}^T \mathbf{w}_i < a$. In matrix notation, we get:

$$f^*(\mathbf{s}) = \sup_{\mathbf{z}} (\mathbf{s}^T \mathbf{W} \mathbf{z} - a \|\mathbf{z}\|_1) = \|\mathbf{s}^T \mathbf{W}\|_{\infty}^a = \|\mathbf{W}^T \mathbf{s}\|_{\infty}^a = \begin{cases} 0 & |\mathbf{s}^T \mathbf{w}_i| \leq a, \forall i \\ +\infty & \text{otherwise} \end{cases}$$

We would reach the same result if we simply noticed that compared to the previous example, the variable \mathbf{s} is now replaced by $\mathbf{W}^T \mathbf{s}$, so that's what goes inside the $\|\cdot\|_{\infty}^a$.

Difference ℓ_1 -norm

Have any suggestions for a new topic? Maybe I can write about it.

the Legendre transform is

$$\sup_{\mathbf{x}} \mathbf{s}^T \mathbf{x} - a \|\mathbf{x} - \mathbf{y}\|_1$$

The transform can be written as the sum of elements of \mathbf{s} and \mathbf{x} as:

$$\sup \sum (s_i x_i - a |x_i - y_i|) = \sup \sum \left(\begin{cases} (s_i - a) x_i + a \cdot y_i & , x_i - y_i \geq 0 \\ (s_i + a) x_i - a \cdot y_i & , x_i - y_i < 0 \end{cases} \right)$$

This way is easy to see that if any element $s_i > a$, then any vector \mathbf{x} with $x_i \rightarrow \infty$ will push the transform to infinity. Likewise, if any element $s_i < -a$, then any vector \mathbf{x} with $x_i \rightarrow -\infty$ will push the transform to ∞ too. For any $-a \leq s_i \leq a$, $(s_i - a) \leq 0$ and $(s_i + a) \geq 0$. Then, under this condition, the expressions $(s_i - a)x_i$ and $(s_i + a)x_i$ are always lower than zero because they always multiply terms with opposite signs. That in turn, means that the supremum is found for $x_i = y_i$, which leads to:

$$\sup_{x_i} (s_i x_i - a |x_i - y_i|) = \begin{cases} (s_i - a)y_i + a \cdot y_i = s_i y_i & x_i - y_i \geq 0 \\ (s_i + a)y_i - a \cdot y_i = s_i y_i & x_i - y_i < 0 \end{cases}$$

In conclusion, the transform is:

$$f^*(\mathbf{s}) = \sup_{\mathbf{x}} (\mathbf{s}^T \mathbf{x} - a \|\mathbf{x} - \mathbf{y}\|_1) = \|\mathbf{s}\|_{\infty}^a = \begin{cases} \mathbf{s}^T \mathbf{y} & |s_i| \leq a, \forall i \\ +\infty & \text{otherwise} \end{cases}$$

When $\mathbf{y} = \mathbf{0}$, the above expression reverts to the ℓ_1 -norm case.

Extra information

- [Explanation of Legendre transformation and pictures](#) supporting it.
- [Legendre-Fenchel transforms in a nutshell](#)

If I helped you in some way, please help me back by liking this website on the bottom of the page or clicking on the link below. It would mean the world to me!

Like { 12

 +1 0

Tweet

Share

11

Have any suggestions for a new topic? Maybe I can write about it. [Let me know](#)