Convex Optimization Discussion - Week 5

University of California, San Diego

February 4, 2022

Outline

1. Conjugate Function
   - Conjugate Function
   - Supporting Hyperplane
   - Conjugate Function Properties
   - Solving Conjugate Function Problems
2. Compressed Sensing
3. Quasiconvex & Quasiconcave Functions
   - Level & Sublevel Sets
   - Quasiconvex & Quasiconcave Functions
   - Solving Quasiconvex & Quasiconcave Function Problems

Conjugate Function

For a function $f : \mathbb{R}^n \to \mathbb{R}$ (that is not necessarily convex), the conjugate $f^* : \mathbb{R}^n \to \mathbb{R}$ is defined as follows:

$$f^*(y) = \sup_{x \,\in\, \mathbf{dom} f} \left(y^T x - f(x)\right)$$

Recall: Let function $g(x, y)$ be convex in y (but not necessarily in x). The function $h(y) = \sup_x g(x, y)$ is a convex function (pointwise maximum-supremum).
Therefore, **the conjugate function is always convex**.

Conjugate Function

The conjugate function represents the negative bias of a line (hyperplane) touching the function at point a certain point.

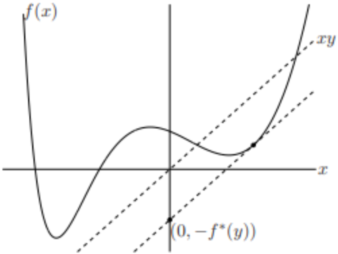


Figure 3.8 A function $f : \mathbf{R} \to \mathbf{R}$, and a value $y \in \mathbf{R}$. The conjugate function $f^*(y)$ is the maximum gap between the linear function yx and $f(x)$, as shown by the dashed line in the figure. If f is differentiable, this occurs at a point x where $f'(x) = y$.

Figure: Boyd & Vandenberghe, Convex Optimization Chapter 3

Supporting Hyperplane

The supporting hyperplane to a function f at point $\bar{x}$ can be derived using the first order condition:

$$\begin{aligned} f(x) &\geq f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) \\ \nabla f(\bar{x})^T \bar{x} - f(\bar{x}) &\geq \nabla f(\bar{x})^T x - f(x) \\ \begin{bmatrix} \nabla f(\bar{x})^T & -1 \end{bmatrix} \begin{bmatrix} \bar{x} \\ f(\bar{x}) \end{bmatrix} &\geq \begin{bmatrix} \nabla f(\bar{x})^T & -1 \end{bmatrix} \begin{bmatrix} x \\ f(x) \end{bmatrix} \\ \bar{b} &\geq \begin{bmatrix} \nabla f(\bar{x})^T & -1 \end{bmatrix} \begin{bmatrix} x \\ f(x) \end{bmatrix} \end{aligned} \tag{1}$$

Supporting Hyperplane

$$f^*(y) = \sup_x \left(y^T x - f(x)\right)$$

$$f^*(y) = \sup_x \left(\begin{bmatrix} y^T & -1 \end{bmatrix} \begin{bmatrix} x \\ f(x) \end{bmatrix}\right)$$

$$y := \nabla f(\bar{x}) \implies f^*(y) = \sup_x \left(\begin{bmatrix} \nabla f(\bar{x})^T & -1 \end{bmatrix} \begin{bmatrix} x \\ f(x) \end{bmatrix}\right) = \bar{b} \qquad (2)$$

Plugging this result back into the previous result:

$$\text{(1) \& (2)} \implies f^*(y) \geq \begin{bmatrix} y^T & -1 \end{bmatrix} \begin{bmatrix} x \\ f(x) \end{bmatrix}$$

$$f^*(y) \geq y^T x - f(x)$$

$$f(x) \geq y^T x - f^*(y)$$

We've found a hyperplane that the function is always greater than or equal to, given that $y = \nabla f(\bar{x})$. The hyperplane $h(x) = y^T x - f^*(y)$ is the **supporting hyperplane**.

Conjugate Function Properties

- Fenchel's Inequality:

$$\forall x \in \mathbf{dom} f,\ y \in \mathbf{dom} f^* : f(x) + f^*(y) \geq x^T y$$

- f is closed (has a closed epigraph) and convex $\implies f^{**} = f$. The functions that have $\mathbf{dom} f = \mathbb{R}^n$ are closed, for example.

Solving Conjugate Function Problems

Let $f^*(y) = \sup_x g(x, y) = \sup_x \left(y^T x - f(x)\right)$. At a given point $\bar{y}$, the conjugate could be one of the following:

- **Finite**: $g(x, \bar{y}) \in \mathbb{R}$ (the good scenario).
- **Infeasible**: $g(x, \bar{y}) \to +\infty$ for **at least one** choice of x. For example, we may find that for $\tilde{x} = [1\ 2]^T$, and $x^{(t)} = t\tilde{x}$; as $t \to \infty$, $g(x, \bar{y}) \to \infty$.
  Intuition: If I can keep making $g(x, \bar{y})$ larger and larger somehow, then all finite values that $g(x, \bar{y})$ can take will eventually pale in comparison.
- **Unbounded Below**: $g(x, \bar{y}) \to -\infty$ for **all** x values. For all sets of $x^{(t)}$ values; as $t \to \infty$, $g(x, \bar{y}) \to -\infty$ (This is rare.)
  Intuition: If there existed any finite solutions anywhere, it would have been preferable to $-\infty$.

Important note: This is just for one value of y. Usually we consider the problem in ranges of y values, and determine if any finite (feasible) solutions exist in those ranges separately.

Solving Conjugate Function Problems

Example: $f(x) = a^T x + b,\ f^*(y) = ?$

$$f^*(y) = \sup_x \left(y^T x - a^T x - b\right) = \sup_x \left((y-a)^T x - b\right)$$

$$y = a \implies f^*(y) = \sup_x(-b) = -b$$

For one index k: $y_k < a_k \implies$ Choose $x^{(t)}$ s.t. $x_i^{(t)} = \begin{cases} -t & i = k \\ 0 & i \neq k \end{cases}$

$$\implies (y-a)^T x - b = (a_k - y_k)t - b$$

$$\implies \lim_{t\to\infty}((a_k - y_k)t - b) \to +\infty$$

We can make the same argument easily for multiple indexes that satisfy $y_i < a_i$. The solution is therefore infeasible, unless $y \succeq a$.

Solving Conjugate Function Problems

$$\text{For one index k: } y_k > a_k \implies \text{Choose } x^{(t)} \text{ s.t. } x_i^{(t)} = \begin{cases} t & i = k \\ 0 & i \neq k \end{cases}$$

$$\implies (y - a)^T x - b = (y_k - a_k)t - b$$

$$\implies \lim_{t \to \infty} ((y_k - a_k)t - b) \to +\infty$$

We can make the same argument easily for multiple indexes that satisfy $y_i > a_i$. The solution is therefore infeasible, unless $y = a$. Putting it all together:

$$f^*(y) = \begin{cases} -b & y = a \\ \infty & y \neq a \end{cases}$$

Solving Conjugate Function Problems

Example (2020 CSE203B QII.3):

$$f(x) = \begin{cases} \frac{1}{2}x^2 & |x| < 1 \\ |x| - \frac{1}{2} & |x| > 1 \end{cases}, \ f^*(y) = ?$$

The absolute value is confusing, rewrite $f(x)$:

$$f(x) = \begin{cases} -x - \frac{1}{2} & x < -1 \\ \frac{1}{2}x^2 & -1 < x < 1 \\ x - \frac{1}{2} & x > 1 \end{cases}$$

Solving Conjugate Function Problems

$$
\begin{aligned}
f^*(y) &= \sup_x \left(y^T x - f(x) \right) \\
&= \sup_x \begin{cases} yx + x + \frac{1}{2} & x < -1 \\ yx - \frac{1}{2}x^2 & -1 < x < 1 \\ yx - x + \frac{1}{2} & x > 1 \end{cases} \\
&= \sup_x \begin{cases} (y+1)x + \frac{1}{2} & x < -1 \\ yx - \frac{1}{2}x^2 & -1 < x < 1 \\ (y-1)x + \frac{1}{2} & x > 1 \end{cases}
\end{aligned}
$$

We can check for cases where $y < -1$ and $y > 1$ to target the spots where we can increase a choice of x in an unbounded fashion.

Solving Conjugate Function Problems

For $y < -1$, we can take $x = -t < -1$. Then:

$$\lim_{t \to \infty} (yx - f(x)) = \lim_{t \to \infty} \left((-y - 1)t + \frac{1}{2} \right) \to +\infty$$

For $y > 1$, we can take $x = t > 1$. Then:

$$\lim_{t \to \infty} (yx - f(x)) = \lim_{t \to \infty} \left((y - 1)t + \frac{1}{2} \right) \to +\infty$$

Therefore, for $y < -1$ and $y > 1$, the conjugate is infeasible. Note that for $-1 \leq y \leq 1$, we cannot argue the same way, since the sign of x is locked to some value in both of the piecewise functions we considered (you can make them go to $-\infty$, but then you still have to check if there's a finite solution elsewhere!)

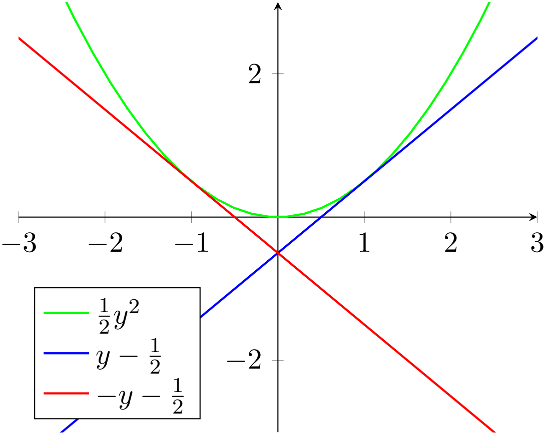
Solving Conjugate Function Problems

Take the derivative to see how each piecewise part behaves when $-1 \leq y \leq 1$:

- $\frac{d\left((y+1)x+\frac{1}{2}\right)}{dx} = y + 1 \geq 0$, with the derivative only being 0 when $y = -1$ (at which point the function is constant.) Since the derivative is positive, we should get as close the solution at the boundary as possible, $x \to -1 \implies \sup_x \left((y+1)x + \frac{1}{2}\right) \to -y - \frac{1}{2}$.
- $\frac{d\left(yx-\frac{1}{2}x^2\right)}{dx} = y - x = 0 \implies \hat{x} = y \implies \sup_x \left(yx - \frac{1}{2}x^2\right) = \frac{1}{2}y^2$.
- $\frac{d\left((y-1)x+\frac{1}{2}\right)}{dx} = y - 1 \leq 0$, with the derivative only being 0 when $y = 1$ (at which point the function is constant.) Since the derivative is negative, we should get as close the solution at the boundary as possible, $x \to 1 \implies \sup_x \left((y-1)x + \frac{1}{2}\right) \to y - \frac{1}{2}$.

Solving Conjugate Function Problems

How to choose maximum? Luckily, we don't have to divide up the y ranges further here, since $\frac{1}{2}y^2$ is always larger than the alternatives:



Solving Conjugate Function Problems

Putting it all together:

$$f^*(y) = \begin{cases} \frac{1}{2}y^2 & -1 \leq y \leq 1 \\ \infty & \text{otherwise} \end{cases}$$

Final comments: In some questions (when you have just a convex quadratic for example) you can just take the derivative to check for optimums and it "works". Still, it pays to know when it's okay to do this. Pay attention to the signs of y, x and infinity values as you're applying the limits, since getting those wrong often leads to completely wrong deductions. Solve the question a bunch of times from scratch if you want to make sure.

Compressed Sensing

A couple of comments that will help with the HW:

- The data is sampled for 5 seconds at 1kHz. This determines the t values, and m (i.e. how many t values to consider).
- The approximation should be

$$y_t \approx \sum_k a_{2k} sin(2\pi f_k t) + a_{2k+1} cos(2\pi f_k t)$$

  Watch out for the k indexes. This is doable without a size error if a has **twice** the size of f.
- Useful: $y = Ax + b \implies b = y - Ax$.
- Good chance to familiarize yourself with CVX(PY).

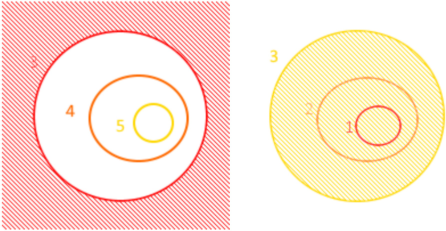
Level & Sublevel Sets

The level set for $f : \mathbb{R}^n \to \mathbb{R}$ at level c is:

$$L_c = \{x \in \mathbf{dom} f \mid f(x) = c\}$$

The sublevel set has a similar definition:

$$S_c = \{x \in \mathbf{dom} f \mid f(x) \leq c\}$$



(a) $\{x \mid f(x) \leq 3\}$, concave.

(b) $\{x \mid g(x) \leq 3\}$, convex.

Figure: Different sublevel sets for different functions.

Quasiconvex & Quasiconcave Functions

A function f is quasiconvex if all its sublevel sets are convex. That is:

$$\forall c: \; \{x \in \mathbf{dom} f \mid f(x) \leq c\} \text{ is convex.}$$

A function f is quasiconcave if all its superlevel sets are convex. That is:

$$\forall c: \; \{x \in \mathbf{dom} f \mid f(x) \geq c\} \text{ is convex, or}$$

$$\forall c: \; \{x \in \mathbf{dom} f \mid -f(x) \leq c\} \text{ is convex.}$$

A function that is both quasiconvex and quasiconcave is said to be quasilinear.

Quasiconvex & Quasiconcave Functions

Important Notes: Convexity and quasiconvexity definitions are related, but often do not imply one another in an intuitive way. A list of things to check:

- f is affine $\implies$ f is quasilinear.
- f is convex and not affine $\implies$ f is quasiconvex and not concave.
- f is concave and not affine $\implies$ f is quasiconcave and not convex.
- f is quasiconvex $\not\Longrightarrow$ f is convex.
- f is quasiconvex $\not\Longrightarrow$ f is not quasiconcave.
- f is quasiconcave $\not\Longrightarrow$ f is concave.
- f is quasiconcave $\not\Longrightarrow$ f is not quasiconvex.

Quasiconvex & Quasiconcave Functions

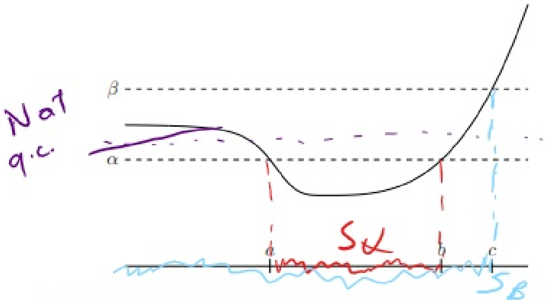
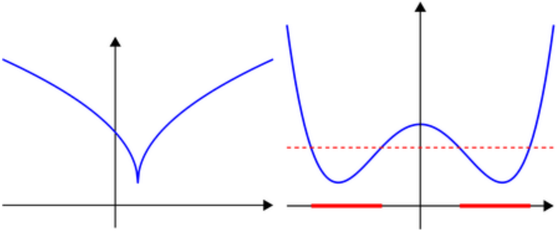


Figure 3.9 A quasiconvex function on $\mathbf{R}$. For each α, the α-sublevel set S_α is convex, *i.e.*, an interval. The sublevel set S_α is the interval $[a, b]$. The sublevel set S_β is the interval $(-\infty, c]$.

Figure: Boyd & Vandenberghe, Convex Optimization Chapter 3

A quasiconvex function does not guarantee that the local minimum equals the global minimum (or that minimum points will satisfy 1st-2nd order conditions etc.), but guarantees that when there's a "dip" in the function, outward from that "dip", the function is non-decreasing.

Quasiconvex & Quasiconcave Functions



(a) Quasiconvex (b) Not quasiconvex

Figure: Wikipedia, Quasiconvex Function

Quasiconvex & Quasiconcave Functions

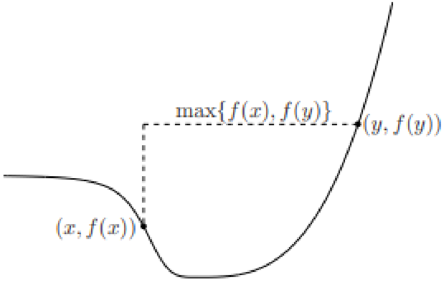


Figure 3.10 A quasiconvex function on **R**. The value of f between x and y is no more than $\max\{f(x), f(y)\}$.

Figure: Boyd & Vandenberghe, Convex Optimization Chapter 3

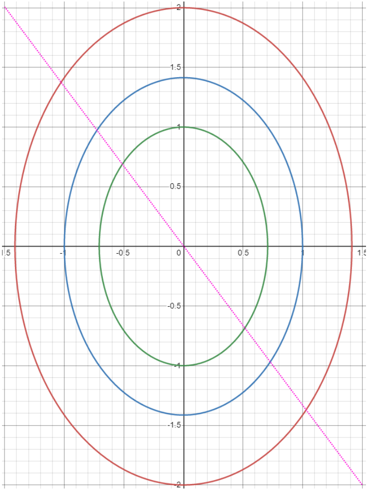
Another definition of a quasiconvex function:

$$\forall x_1, x_2 \in \mathbf{dom} f,\ \theta \in [0,1] : f\left(\theta x_1 + (1-\theta)x_2\right) \leq \max\{f(x_1), f(x_2)\}$$

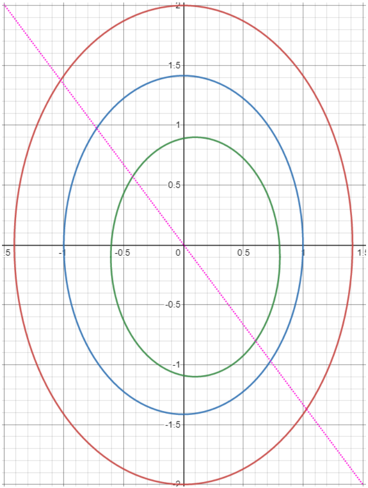
Solving Quasiconvex & Quasiconcave Function Problems

Level Set Diagram Questions: You may be asked to comment on convexity/quasiconvexity given the level sets of a function. It is easy to comment on sublevel sets, since the level sets are given (you shouldn't need consider anything else for the quasiconvex - quasiconcave possibilities.) If you figure out that the funciton is quasiconvex, there's a chance it may be convex as well (same for quasiconcave and concave), but there's a catch!

Solving Quasiconvex & Quasiconcave Function Problems



(a) Convex level sets



(b) Non-convex level sets

Solving Quasiconvex & Quasiconcave Function Problems

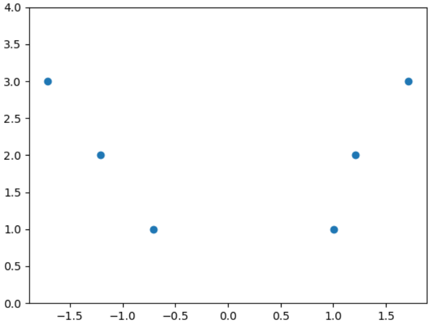


Figure: Cross-section of the non-convex function at (b)

Takeaway: When there's an obvious imbalance in cross-sections in the image, it should be considered as you're determining convexity (but not quasiconvexity, as that is more general.)

Solving Quasiconvex & Quasiconcave Function Problems

Algebraic Questions: Some important strategies:

- Check for convexity/concavity. If you can prove convexity, you get quasiconvexity for free (and the same for concavity - quasiconcavity). For example, check for PSD/NSD for the Hessian. A trick: if the Hessian is expressible in terms of an outer product, you get convexity:

$$x^T H x = x^T \left(uu^T\right) x = \left(u^T x\right)^T \left(u^T x\right) = ||u^T x||_2^2 \geq 0$$

- A good way to check for quasiconvex - quasiconcave is to plot the function. If you can visualize the sublevel sets, it's easier to determine if they'll always be convex or not. **Keep in mind that points outside the domain of f do not violate quasiconvexity - quasiconcavity**.
- For continuous functions $f : \mathbb{R} \to \mathbb{R}$: Let $c \in \mathbb{R} \cup \{-\infty, +\infty\}$. f is quasiconvex $\iff x \leq c$ the function is non-increasing and $x > c$ it is non-decreasing (for one such c only).

Solving Quasiconvex & Quasiconcave Function Problems

Example (Boyd & V. 3.31): $f(x_1, x_2) = x_1 x_2,\ \mathbf{dom} f = \mathbb{R}^2_+$
convex/concave/quasiconvex/quasiconcave?
Check the Hessian:

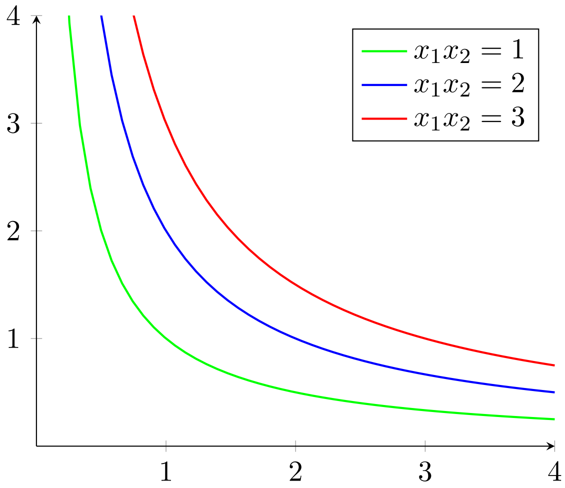
$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$u := \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies u^T \nabla^2 f(x_1, x_2) u = 2 > 0$$

$$v := \begin{bmatrix} -1 \\ 1 \end{bmatrix} \implies v^T \nabla^2 f(x_1, x_2) v = -2 < 0$$

Therefore, f is not convex and not concave.

Solving Quasiconvex & Quasiconcave Function Problems



Plotting the function, we can see that the superlevel sets are convex and sublevel sets are not. Therefore, f is quasiconcave and not quasiconvex.