Lecture 21: Constrained Optimization: Motivation and Intuition

Charles A. Bouman

Purdue University

Wednesday, October 14, 2026

Outline

1. Splitting the variable
   - Why split; the penalty $\frac{a}{2}||x - v||^2$ and the correction term u
   - The split-variable augmented Lagrangian and ADMM algorithms
2. Example: positivity constraints
   - The extended function $\tilde{g}$; the v update is a clip
3. What justifies this?

Reading: Section 9.1.

Today's Big Ideas

1. Two equal copies of one variable let each term use its own method.
2. Penalty plus correction u meets the constraint; a simple rule sets u.
3. Minimizing over the two variables in alternation is ADMM.
4. Positivity adds a term (zero inside, infinite outside); its update is a clip.

Split the variable

Start from the MAP problem with two terms,

$$\hat{x} = \underset{x \in \mathbb{R}^N}{\arg\min} \left\{ f(x) + g(x) \right\} \ ,$$

where $f(x)$ is the negative log likelihood and $g(x)$ is the negative log prior probability.

Split the variable into two, x and v, with the constraint $x = v$:

$$(\hat{x}, \hat{v}) = \underset{\substack{(x,v) \\ x=v}}{\arg\min} \left\{ f(x) + g(v) \right\} \ .$$

- The constraint $\hat{x} = \hat{v}$ gives the original solution.
- This looks pointless. Why do it?

Why split the variable

- Often f and g differ so much that no one method suits both.
- After the split, each term is minimized by its own method.
- x and v move on separate paths and converge to one solution.

The example today is positivity. Total variation and plug-and-play priors come in later lectures.

First, solve the constrained problem and see where it leads.

A penalty enforces the constraint, but not exactly

Add a term that penalizes the difference between x and v:

$$(\hat{x}, \hat{v}) = \underset{(x,v)}{\arg\min} \left\{ f(x) + g(v) + \frac{a}{2} ||x - v||^2 \right\} \ ,$$

where $a > 0$ controls the gain of the penalty.

- Larger a enforces $x = v$ more strongly: $\hat{x} \approx \hat{v}$.
- No value of a makes $\hat{x} = \hat{v}$ exactly.
- So the solution is not exactly the solution of the original problem.

The correction term u

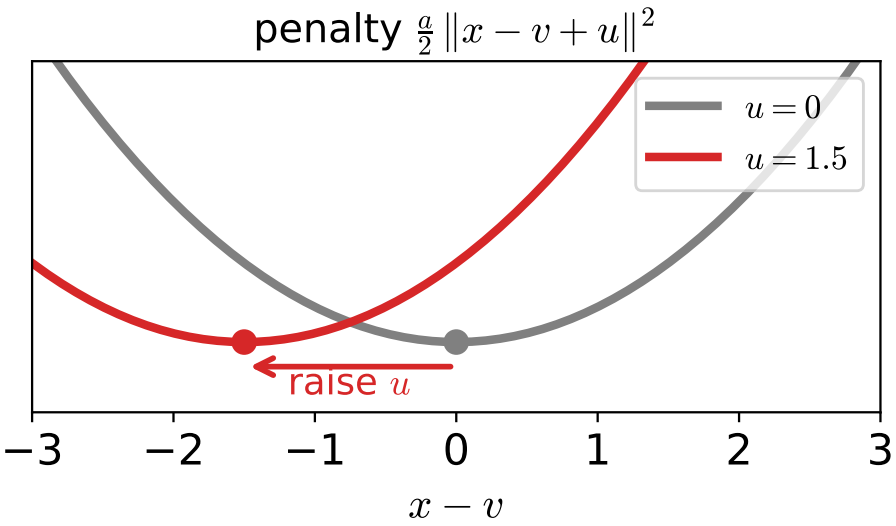
Add a term u inside the penalty that pushes the solution onto the constraint:

$$(\hat{x}, \hat{v}) = \underset{(x,v)}{\arg\min} \left\{ f(x) + g(v) + \frac{a}{2} ||x - v + u||^2 \right\} ,$$

where u must be chosen so that $\hat{x} = \hat{v}$.

The right value of u comes from a simple rule.

How to pick u



The penalty $\frac{a}{2}||x-v+u||^2$ is smallest when $x-v=-u$.

So the penalty pulls $x-v$ toward $-u$.

- Raising u pulls $x-v$ down.
- Lowering u pulls $x-v$ up.

u sets a cost on the difference $(x-v)$.

Rule: adjust u until $\hat{x}-\hat{v}=0$.

- If $\hat{x}-\hat{v}>0$, then raise u.
- If $\hat{x}-\hat{v}<0$, then lower u.
- The simplest update: $u \leftarrow u+(\hat{x}-\hat{v})$.

The split-variable augmented Lagrangian algorithm

Split-Variable Augmented Lagrangian Algorithm

Initialize $u \leftarrow \mathbf{0}$ and $\hat{v} \leftarrow \mathbf{0}$

For K iterations $\{$

$$(\hat{x}, \hat{v}) \leftarrow \underset{(x,v)}{\arg\min} \left\{ f(x) + g(v) + \frac{a}{2} ||x - v + u||^2 \right\}$$

$$u \leftarrow u + (\hat{x} - \hat{v})$$

$\}$

- The first step minimizes with the current u.
- The second step adjusts the cost u until $\hat{x} = \hat{v}$.

The split-variable ADMM algorithm

The joint minimization over (x, v) is hard. Minimize over x and v in alternation instead.

Split-Variable ADMM Algorithm

Initialize $u \leftarrow \mathbf{0}$ and $\hat{v} \leftarrow \mathbf{0}$

For K iterations $\{$

$$\hat{x} \leftarrow \arg\min_x \left\{ f(x) + \frac{a}{2} ||x - \hat{v} + u||^2 \right\}$$

$$\hat{v} \leftarrow \arg\min_v \left\{ g(v) + \frac{a}{2} ||\hat{x} - v + u||^2 \right\}$$

$$u \leftarrow u + (\hat{x} - \hat{v})$$

$\}$

This is the **alternating direction method of multipliers**, or **ADMM**. For convex problems, it converges to the solution.

Example: positivity constraints

- Positivity constraints are common in imaging.
- X is often energy or absorption, so it must be positive.
- The MAP estimate must then use the constraint $X \in \mathbb{R}^{+N}$.

The two-term form $f(x) + g(x)$ can enforce this constraint with a special choice of g.

Positivity as an unconstrained problem

The constrained problem is

$$\hat{x} = \underset{x \in \mathbb{R}^{+N}}{\arg\min}\, f(x) \ .$$

It equals the unconstrained problem

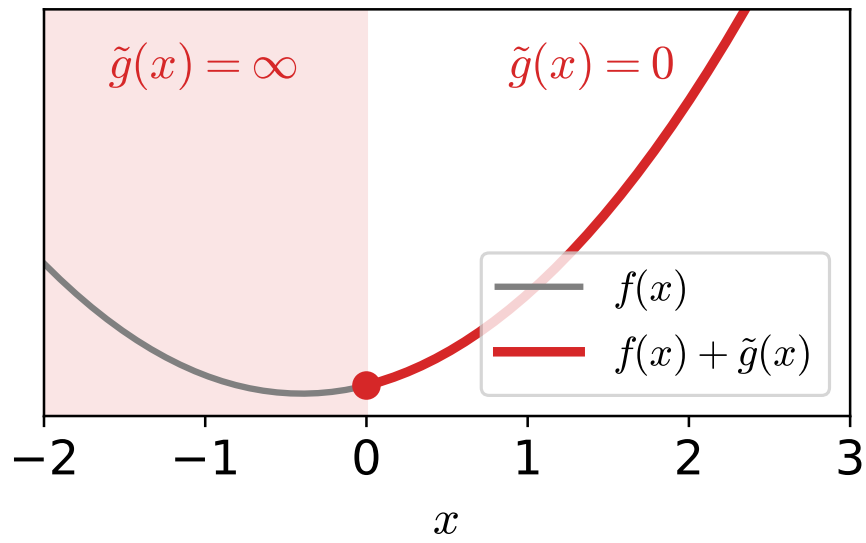
$$\hat{x} = \underset{x \in \mathbb{R}^{N}}{\arg\min} \left\{ f(x) + \tilde{g}(x) \right\} \ ,$$

where

$$\tilde{g}(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R}^{+N} \\ \infty & \text{if } x \notin \mathbb{R}^{+N} \end{cases} \ .$$

The function $\tilde{g}$

- $\tilde{g}$ is an **extended function**: its values are in $\mathbb{R} \cup \{\infty\}$.
- $\tilde{g}$ is **convex**, with the convention $\infty + a = \infty$.



Define: an extended function h is

- **proper** if $h(x) < \infty$ for some x, and
- **closed** if every sublevel set $\{x : h(x) \leq \alpha\}$ is a closed set.

$\tilde{g}$ is proper, since $\tilde{g}(0) = 0$, and closed, since every sublevel set is $\mathbb{R}^{+N}$ or empty.

So $\tilde{g}$ is a **proper closed convex function**. Theorems about extended convex functions assume all three properties.

Positivity: the v update is a clip

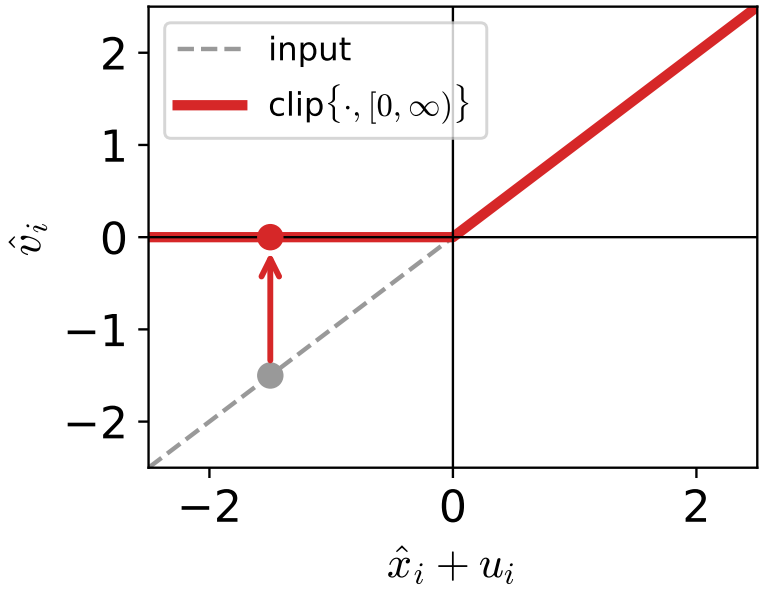
With $g = \tilde{g}$, the v update has a closed form:

$$\begin{aligned} \hat{v} &= \arg\min_v \left\{ \tilde{g}(v) + \tfrac{a}{2} ||\hat{x} + u - v||^2 \right\} \\ &= \arg\min_{v \in \mathbb{R}^{+N}} ||\hat{x} + u - v||^2 \\ &= \text{clip}\left\{ \hat{x} + u, \mathbb{R}^{+N} \right\} \ . \end{aligned}$$

Define $\text{clip}\{x, \mathcal{A}\}$ as the closest point to x in the closed convex set $\mathcal{A}$.

For $\mathcal{A} = \mathbb{R}^{+N}$, each negative component is set to zero:

$$\hat{v}_i = \max\{\hat{x}_i + u_i, 0\} \ .$$



Split-variable ADMM for a positivity constraint

Split-Variable ADMM for Positivity Constraint

Initialize $u \leftarrow \mathbf{0}$ and $\hat{v} \leftarrow \mathbf{0}$

For K iterations $\{$

$$\begin{aligned} \hat{x} &\leftarrow \arg\min_x \left\{ f(x) + \frac{a}{2} ||x - (\hat{v} - u)||^2 \right\} \\ \hat{v} &\leftarrow \text{clip}\left\{ \hat{x} + u, \mathbb{R}^{+N} \right\} \\ u &\leftarrow u + (\hat{x} - \hat{v}) \end{aligned}$$

$\}$

- The x update minimizes f plus a quadratic; any solver for f works.
- The v update is a clip.
- This enforces positivity in any convex optimization problem.

What justifies this?

So far, the approach is a heuristic.

- Split the variable, penalize the difference, correct with u.
- Adjust u by the error $\hat{x} - \hat{v}$.
- Minimize over x and v in alternation.

Open questions.

- Does u converge? To what?
- Does the solution meet the constraint $\hat{x} = \hat{v}$?
- Does the alternating version converge to the same answer?

The answers need a formal theory: Lagrange multipliers, duality, and the augmented Lagrangian. That is the next lecture.

Summary

1. Splitting $\arg\min_x\{f(x)+g(x)\}$ into $\arg\min_{(x,v),\,x=v}\{f(x)+g(v)\}$ lets each term be minimized separately.

2. The penalty and correction give $\arg\min_{(x,v)}\{f(x)+g(v)+\frac{a}{2}||x-v+u||^2\}$, and the update $u \leftarrow u+(\hat{x}-\hat{v})$ drives $\hat{x}=\hat{v}$.

3. ADMM minimizes over x and v in alternation.

4. Positivity is an added extended function, $\tilde{g}(x)=0$ on $\mathbb{R}^{+N}$ and ∞ elsewhere, and its v update is $\hat{v}=\text{clip}\{\hat{x}+u,\mathbb{R}^{+N}\}$.

Next lecture: Lagrange multipliers, duality, and the augmented Lagrangian, which justify this approach.