Lecture 19: Surrogate Properties, and Building the Quadratic Surrogate

Chapter 8.2–8.3

Charles A. Bouman

Purdue University

Wednesday, October 7, 2026

Outline

1. **Five surrogate properties**: behavior, and how they are built
2. **Applying to the MAP cost**: only the prior needs a surrogate
3. **Quadratic surrogate**: maximum curvature and symmetric bound
4. **Coefficients**: when the bound is valid, and for which potentials

Reading: Chapter 8.2–8.3, *Foundations of Computational Imaging*.

Today's Big Ideas

- The surrogate properties let us build the surrogate one term at a time.
- The data term is quadratic, so only the prior needs a surrogate
- Maximum curvature is too conservative; symmetric bound is tight
- Huber and QGGMRF work; total variation and GGMRF do not

Three properties: how the algorithm behaves

A surrogate function $Q(x; x')$ for minimizing $f(x)$ has five properties. The first three describe the majorization minimization algorithm.

Monotonicity

Any x that lowers Q below its value at x' also lowers f: $Q(x; x') < Q(x'; x') \Rightarrow f(x) < f(x')$. We do not need the exact minimizer of Q.

Gradient

At the point of approximation, Q and f have the same gradient, $\nabla f(x') = \nabla Q(x; x')|_{x=x'}$. The two functions are tangent there.

Fixed point

If x^* minimizes $Q(x; x^*)$, then $\nabla f(x^*) = 0$. For a convex f this is a local minimum; for a strictly convex f it is the global minimum.

Two properties: how surrogates are built

The last two properties let us build a surrogate for a large cost function out of surrogates for its parts.

Additivity: a sum of surrogates is a surrogate of the sum

If $Q_k(x; x')$ is a surrogate for $f_k(x)$ for each k, then $\sum_k Q_k(x; x')$ is a surrogate for $\sum_k f_k(x)$.

Composition: composing with a function preserves the property

If $Q(x; x')$ is a surrogate for $f(x)$ and $g : \mathcal{A} \to \mathcal{A}$, then $Q(g(x); g(x'))$ is a surrogate for $f(g(x))$.

Together these mean we can design a surrogate for one pairwise term of the prior and then add the terms up to get a surrogate for the whole cost function.

Applying the properties to the MAP cost function

The MAP cost function is

$$f(x) = \frac{1}{2}||y - Ax||_\Lambda^2 + \sum_{\{s,r\}\in\mathcal{P}} b_{s,r}\rho(x_s - x_r) \ .$$

- The data term is quadratic; the **prior term** needs a quadratic surrogate
- By additivity and composition, a potential surrogate gives a prior one

With a quadratic surrogate potential $\rho(\Delta; \Delta')$, the surrogate MAP cost function is

$$Q(x; x') = \frac{1}{2}||y - Ax||_\Lambda^2 + \sum_{\{s,r\}\in\mathcal{P}} b_{s,r}\rho(x_s - x_r; x'_s - x'_r) \ ,$$

a quadratic function of x with the Chapter 5 form. The remaining goal is a quadratic surrogate for the potential, $\rho(\Delta; \Delta')$.

The maximum curvature method

The **maximum curvature** method uses a surrogate potential of the form

$$\rho(\Delta;\Delta') = a_1\Delta + \frac{a_2}{2}\left(\Delta - \Delta'\right)^2 \ .$$

Matching gradients at $\Delta = \Delta'$ gives $a_1 = \rho'(\Delta')$. To make the surrogate an upper bound, set a_2 to the largest second derivative of the potential,

$$a_2 = c_{max} \triangleq \max_{\Delta\in\mathbb{R}} \rho''(\Delta) \ .$$

Here c_{max} is fixed for a given potential, so only a_1 is recomputed as Δ' changes.

Disadvantage

The curvature $a_2 = c_{max}$ is **too conservative**, especially when the second derivative is large near zero. If the second derivative is infinite, there is no maximum and **the method cannot be used**.

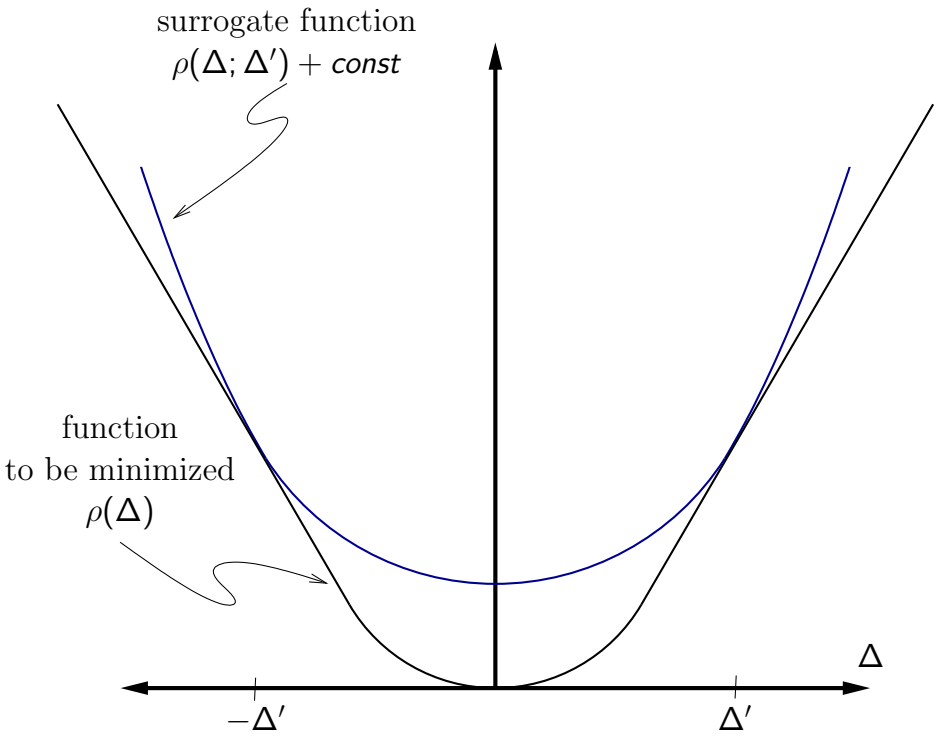
The symmetric bound method

The **symmetric bound** method bounds the potential by a **symmetric quadratic**,

$$\rho(\Delta;\Delta') = \frac{a_2}{2}\,\Delta^2 \ ,$$

where a_2 depends on Δ'.

- Works for all of Section 6.3
- Gives a much tighter bound than maximum curvature, so the optimization converges faster.



The curvature a_2 is chosen so the quadratic touches the potential at two points symmetric about zero. This gives a much tighter surrogate than a single maximum curvature.

Determining the coefficient a_2

Matching gradients of $\rho(\Delta)$ and $\rho(\Delta;\Delta')$ at $\Delta = \Delta'$ gives

$$a_2 = \frac{\rho'(\Delta')}{\Delta'} \ ,$$

so the symmetric bound surrogate potential is

$$\rho(\Delta;\Delta') = \begin{cases} \dfrac{\rho'(\Delta')}{2\Delta'}\Delta^2 & \text{if } \Delta' \neq 0 \ , \\ \dfrac{\rho''(0)}{2}\Delta^2 & \text{if } \Delta' = 0 \ . \end{cases}$$

At $\Delta' = 0$ we use the limiting value $a_2 = \rho''(0)$.

When the symmetric bound is a valid upper bound

We must still check the upper bound condition $\rho(\Delta) \leq \rho(\Delta; \Delta')$ from the definition of a surrogate function.

Sufficient condition

The condition holds for any symmetric potential whose influence function $\rho'(\Delta)$ is

- **monotone increasing near** $\Delta = 0$, and
- **concave for** $\Delta > 0$.

Most practical potentials satisfy this, since it means the influence of a neighbor decreases as the difference between the two pixels grows.

Coefficients for the symmetric bound (1 of 2)

Name	$\rho(\Delta)$	$\frac{\rho'(\Delta')}{2\Delta'}$	$\frac{\rho''(0)}{2}$
Quadratic	$\frac{\Delta^2}{2\sigma_x^2}$	$\frac{1}{2\sigma_x^2}$	$\frac{1}{2\sigma_x^2}$
Total Variation	$\frac{\lvert\Delta\rvert}{\sigma_x}$	$\frac{1}{2\sigma_x\lvert\Delta'\rvert}$	∞
Huber	$\begin{cases} \frac{\Delta^2}{2\sigma_x^2} & \text{for } \lvert\Delta\rvert < T\sigma_x \ , \\ T\left\lvert\frac{\Delta}{\sigma_x}\right\rvert - \frac{T^2}{2} & \text{for } \lvert\Delta\rvert \geq T\sigma_x \end{cases}$	$\frac{\text{clip}\{\Delta', [-T\sigma_x, T\sigma_x]\}}{2\sigma_x^2\Delta'}$	$\frac{1}{2\sigma_x^2}$
log cosh	$T^2 \log\cosh\left(\frac{\Delta}{T\sigma_x}\right)$	$\left\lvert\frac{T\tanh\left(\frac{\Delta'}{T\sigma_x}\right)}{2\sigma_x\Delta'}\right\rvert$	$\frac{1}{2\sigma_x^2}$

The middle column is the curvature $a_2/2$ at a general Δ'; the right column is its limit at $\Delta' = 0$.

Coefficients for the symmetric bound (2 of 2)

Name	$\rho(\Delta)$	$\frac{\rho'(\Delta')}{2\Delta'}$	$\frac{\rho''(0)}{2}$
GGMRF	$\frac{\lvert\Delta\rvert^p}{p\sigma_x^p}$	$\frac{\lvert\Delta'\rvert^{p-2}}{2\sigma_x^p}$	∞ for $1 \leq p < 2$
QGGMRF ($p \leq q$)	$\frac{\lvert\Delta\rvert^p}{p\sigma_x^p}\left(\frac{\left\lvert\frac{\Delta}{T\sigma_x}\right\rvert^{q-p}}{1+\left\lvert\frac{\Delta}{T\sigma_x}\right\rvert^{q-p}}\right)$	$\frac{\lvert\Delta'\rvert^{p-2}}{2\sigma_x^p}\frac{\left\lvert\frac{\Delta'}{T\sigma_x}\right\rvert^{q-p}\left(\frac{q}{p}+\left\lvert\frac{\Delta'}{T\sigma_x}\right\rvert^{q-p}\right)}{\left(1+\left\lvert\frac{\Delta'}{T\sigma_x}\right\rvert^{q-p}\right)^2}$	$\frac{1}{p\sigma_x^2 T^{2-p}}$ for $q=2$

Each row corresponds to a potential function from the table of Section 6.3.

Bounded and unbounded second derivatives at zero

- Some entries for $\rho''(0)/2$ in the tables are **unbounded**.
- Total variation and GGMRF with $1 \leq p < 2$: infinite $\rho''(0)$
- **This breaks the symmetric bound surrogate at $\Delta' = 0$.**

Potentials compatible with surrogate optimization

The Huber and QGGMRF potentials have a **bounded** second derivative at zero, so both are compatible with surrogate optimization.

Chapter 9 gives algorithms for the potentials with an unbounded second derivative at zero, such as total variation.

Summary

- Three properties describe the algorithm; two build surrogates
- Only the prior needs a surrogate; the data term is quadratic
- Maximum curvature is simple but too conservative.
- The symmetric bound is tighter and general, with weight $\rho'(\Delta)/\Delta$.
- Works for Huber and QGGMRF; not for total variation or GGMRF

Reading: Chapter 8.2–8.3. **Next lecture:** the reweighted quadratic update, the surrogate line search, and ICD.