Linearization of the Biot Savart Law

Sara Loupot
David Fuentes

May 23, 2015

1 Introduction

The Biot Savart Law describes the magnetic field as a function of distance from a magnetic dipole source. This relationship is used in magnetic relaxometry to relate the signal received by the SQUID pickup coils to the location of the bound nanoparticles. In order to create an image, the magnetic inverse problem must be solved. This problem is ill-posed: there are more unknown parameters than known. To employ state of the art methods to attempt to solve this problem, the non-linear Biot Savart relationship must be made into a linear form, Ax=b. Under some assumptions, this can be done.

2 The Biot Savart Law

The Biot Savart Law states

$$\vec{B}(\vec{r},\vec{m}) = \frac{\mu_0}{4\pi}\left[\frac{3(\vec{m}*\vec{r})\vec{r}}{r^5} - \frac{\vec{m}}{r^3}\right] \tag{1}$$

Where $\vec{B}$ is the magnetic field at a point $\vec{r}$ from a dipole $\vec{m}$. Immediately after the field is turned off, the moment vector can be approximated as only having a non-zero longitudinal component, $\vec{m} \approx m\hat{z}$, which simplifies the dot product:

$$\vec{m}*\vec{r} = m_i r_i = m r_z$$

and the second term:

$$\frac{\vec{m}}{r^3} = \frac{m}{r^3}\hat{z}$$

The SQUIDs only detect $\vec{B}$ parallel to their axis. If we assume they are arranged parallel to the z axis then

$$\vec{B}(\vec{r},\vec{m}) = (\vec{B}(\vec{r}))_z = \frac{\mu_0}{4\pi}\left[\frac{3mr_z}{r^5}r_z - \frac{m}{r^3}\right]\hat{z} \tag{2}$$

which allows us to factor out m:

$$\vec{B}(\vec{r},\vec{m}) = \frac{\mu_0}{4\pi}m[\frac{r_z^2}{r^5} - \frac{1}{r^3}]\hat{z} \tag{3}$$

Now, B is separable into independent functions that depend on the moment and the distance.

$$\vec{B} = f(m)g(\vec{r}) \tag{4}$$

where

$$f(m) = \frac{\mu_0}{4\pi}m \tag{5}$$

and

$$g(\vec{r}) = \frac{r_z^2}{r^5} - \frac{1}{r^3} \tag{6}$$

2.1 Discretization

We will assume a Galerkin discretization of our solution field, $m_z(r)$.

$$m(r) = \sum_j m_j\phi_j(r) \qquad \phi_j = \begin{cases} 1, & r \in \Omega_j \\ 0, & r \notin \Omega_j \end{cases}$$

Here the imaging domain, Ω, is decomposed into disjoint pixels, Ω_j

$$\Omega = \cup_j \Omega_j \qquad \Omega_i \cap \Omega_j = \emptyset$$

Finally, the field detected by sensor i is a linear combination of the field produced by each source m_j located at $\vec{r}_{ij}$ from the sensor.

$$\vec{B}_i(m) = \frac{\mu_0}{4\pi}\left(\sum_j m_j \phi_j(\vec{r}_{ij})\right) \underbrace{\left[\frac{r_{z_{ij}}^2}{r_{ij}^5} - \frac{1}{r_{ij}^3}\right]}_{g(\vec{r}_{ij})} \hat{z}$$

We then assume a discretized field of view with j pixels, each with their own moment m_j a distance $\vec{r}_{ij}$ from sensor i. To construct the linear problem, define a $n \times 1$ vector $\vec{B}$ of the field detected by sensor i where n is the number of sensors, a $m \times 1$ vector $\vec{x}$ of the source from pixel j where m is the number of pixels, and a $m \times n$ matrix A of $g(\vec{r}_{ij})$ where $\vec{r}_{ij}$ is the vector from pixel j to sensor i, as defined in equation 6.

$$\begin{bmatrix} B_1 \\ \vdots \\ B_i \\ \vdots \\ B_n \end{bmatrix} = \begin{bmatrix} g(\vec{r}_{1,1}) & g(\vec{r}_{1,2}) & \dots & g(\vec{r}_{1,m}) \\ g(\vec{r}_{2,1}) & g(\vec{r}_{2,2}) & \dots & g(\vec{r}_{2,m}) \\ \vdots & \vdots & \ddots & \vdots \\ g(\vec{r}_{n,1}) & g(\vec{r}_{n,2}) & \dots & g(\vec{r}_{n,m}) \end{bmatrix} \begin{bmatrix} m_1 \\ \vdots \\ m_j \\ \vdots \\ m_m \end{bmatrix} \qquad \Leftrightarrow b = A\, x \tag{7}$$

Since the location of the detectors with respect to the field of view grid points is defined, $\vec{x}$ is the only unknown.

Columns of the matrix, A, may be interpreted as 'distance' vectors from magnetic source j to each sensor.

$$A = [\vec{g}_1 \quad \vec{g}_2 \quad \dots \quad \vec{g}_m] \qquad \vec{g}_j = \begin{bmatrix} g(\vec{r}_{1,j}) \\ g(\vec{r}_{2,j}) \\ \vdots \\ g(\vec{r}_{n,j}) \end{bmatrix}$$

2.2 Bias Correction

The bias correction of [1] is of the form

$$\vec{B}_i(m) = \frac{\mu_0}{4\pi}\left(\sum_j \frac{m_j}{\sqrt{\sum_{k=1}^n g^2(\vec{r}_{kj})}} \phi_j(\vec{r}_{ij})\right) \underbrace{\left[\frac{r_{z_{ij}}^2}{r_{ij}^5} - \frac{1}{r_{ij}^3}\right]}_{g(\vec{r}_{ij})} \hat{z}$$

The artificial term, $\sqrt{\sum_{k=1}^n g^2(\vec{r}_{kj})}$, represents the average distance of magnetic source j to the sensors and allows numerical techniques more preferentially weight terms further from the sensor. *@saraloupot Is the physics is preserved ? NEED DIMENSIONAL ANALYSIS. units of m must be changed to match the b-field.*

$$\vec{B}_i\left[\frac{N}{Am}\right] = \frac{\mu_0}{4\pi}\left[\frac{N}{A^2}\right]\left(\sum_j \frac{m_j[Am^2]}{\sqrt{\sum_{k=1}^n g^2(\vec{r}_{kj})}} \phi_j(\vec{r}_{ij})\right) g(\vec{r}_{ij}) \left[\frac{1}{m^3}\right] \tag{8}$$

$$\left[\frac{N}{Am}\right] = \left[\frac{N}{A^2}\right][Am^2]\left[\frac{1}{m^3}\right] \tag{9}$$

$$r \quad \uparrow \quad \Rightarrow \quad g(r) \quad \downarrow \quad \Rightarrow \quad \frac{1}{g(r)} \quad \uparrow$$

The augmented equations is now of the form

$$b = \hat{A}\, \hat{x}$$

where columns of the augmented system matrix are unit normalized

$$\hat{A} = \left[\frac{\vec{g}_1}{\|\vec{g}_1\|} \quad \frac{\vec{g}_2}{\|\vec{g}_2\|} \quad \dots \quad \frac{\vec{g}_m}{\|\vec{g}_m\|}\right] \qquad \vec{g}_j = \begin{bmatrix} g(\vec{r}_{1,j}) \\ g(\vec{r}_{2,j}) \\ \vdots \\ g(\vec{r}_{n,j}) \end{bmatrix}$$

3 Linear Inverse solve

It would seem that the solution is trivial, with a simple linear inverse solve $x = A^{-1}b$. However, since $n << m$ the problem is still ill-posed. There are an infinite number of exact solutions. A simple inverse solution results in a solution of a source under each detector proportional to the strength of the signal received from that detector. Clearly this is a viable solution, but it is unlikely to be the true solution.

4 l_1 norm minimization

In the MRX application, we can assume that the true solution will be sparse: only a few voxels will have a non-zero contribution to the net moment. To increase the chances of converging to the true solution, we can attempt to find the minimum number of sources that still solve Equation 7 exactly. The l_p norm can be used to define sparsity with parameter p.

$$l_p(\vec{x}) = \left[\sum_i x_i^p\right]^{1/p} \tag{10}$$

When p=2, the l_p norm is the magnitude of the vector $\vec{x}$. When p=1, l_p is the sum of the components of $\vec{x}$. When p=0, l_p is the number of non-zero entities of $\vec{x}$. Ideally, to find the sparsest solution that satisfies the equality constraint, it would be best to minimize the l_p norm. Since the l_0 norm is not convex, it is difficult to compute explicitly. Instead, we must maximize sparsity by minimizing the L1 norm of x while applying the condition that $Ax = b$.

$$\begin{array}{ll} \min & \sum_i |x_i| \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \tag{11}$$

We will solve this as a sequence of unconstrained Quadratic Penalty problems with a non-smooth objective function

$$\min_x \|\Phi x\|_1 + \frac{1}{2\mu}\|Ax - b\|_2^2$$

For large μ, the penalty function does not accurately enforce the constraint. As we have seen, typical approaches increasingly impose the constraint by letting $\mu \to 0$. An alternative is to apply the Augmented Lagrangian approach or the so-called Bregman iteration [Goldstein and Osher, 2009] where in the penalty term constants is fixed and we more accurately solve the constraint at each iteration. These approach introduce an auxillary variable $w \in \mathbb{R}^n$ and enforce the transformation to this variable.

$$\min_x \frac{1}{2}\|Ax - b\|_2^2 + \tau\|w\|_1 \qquad \text{such that} \quad \Phi x = w$$

This may be interpreted in our previous notation as

$$f(\hat{x}) = f(x, w) = \frac{1}{2}\|Ax - b\|_2^2 + \tau\|w\|_1 \qquad \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \end{bmatrix} = \Phi x - w$$

and the Augmented Lagrangian may be written as

$$\mathcal{L}_a(\underbrace{x, w}_{\hat{x}}, \lambda; \mu) = \underbrace{\frac{1}{2}\|Ax - b\|_2^2 + \tau\|w\|_1}_{f(\hat{x})} - \underbrace{(\lambda, \Phi x - w)}_{\sum_i \lambda_i c_i(\hat{x})} + \underbrace{\frac{1}{2\mu}\|(\Phi x - w)\|_2^2}_{\frac{1}{2\mu}\sum_i c_i^2(\hat{x})}$$

within the scope of Algorithm **??**, given λ^k and μ^k, the subproblem is still formidable and contains L1 and L2 terms

$$\min_{x,w} \mathcal{L}_a(\cdot, \lambda^k; \mu^k) = \min_{x,w} \frac{1}{2}\|Ax - b\|_2^2 + \tau\|w\|_1 - (\lambda^k, \Phi x - w) + \frac{1}{2\mu}\|(\Phi x - w)\|_2^2$$

The Lagrangian may be written in an equivalent form, using the linearity of the inner product

$$\frac{1}{2c}(a - cb, a - cb) = \frac{1}{2c}\left[(a - cb, a) - (a - cb, cb)\right] = \frac{1}{2c}\|a\|^2 - \frac{1}{2c}c(b, a) - \frac{1}{2c}c(a, b) + \frac{c^2}{2c}\|b\|^2 = \frac{1}{2c}\|a\|^2 - (a, b) + \frac{c}{2}\|b\|^2$$

$$\Rightarrow \qquad \frac{1}{2c}\|a - cb\|^2 - \frac{c}{2}\|b\|^2 = \frac{1}{2c}\|a\|^2 - (a, b)$$

and minimizing with respect to (x, w), $\frac{\mu}{2}\|\lambda^k\|_2^2$ may be considered a constant

$$\min_{x,w} \mathcal{L}_a(\cdot, \lambda^k; \mu^k) = \min_{x,w} \frac{1}{2}\|Ax-b\|_2^2+\tau\|w\|_1+\frac{1}{2\mu}\|\Phi x-w-\mu\lambda^k\|_2^2-\frac{\mu}{2}\|\lambda^k\|_2^2 \quad \Leftrightarrow \quad \min_{x,w} \frac{1}{2}\|Ax-b\|_2^2+\tau\|w\|_1+\frac{1}{2\mu}\|\Phi x-w-\mu\lambda^k\|_2^2$$

The approach becomes tractable if we break the subproblem up into an L1 subproblem and L2 subproblem using an alternating direction or coordinate descent techinque.

$$\underbrace{\text{L2 subproblem}}_{w^k,\lambda^k \text{ fixed}} \qquad \min_x \frac{1}{2}\|Ax-b\|_2^2 + \frac{1}{2\mu}\|\Phi x - w^k - \mu\lambda^k\|_2^2$$

$$\underbrace{\text{L1 subproblem}}_{x^k,\lambda^k \text{ fixed}} \qquad \min_w \tau\|w\|_1 + \frac{1}{2\mu}\|\Phi x^k - w - \mu\lambda^k\|_2^2$$

A summary of the Augmented Lagrangian Algorithm in this context is presented in Algorithm 1. Notice that we have converted the constrained problem into two unconstrained problems.

Algorithm 1 Augmented Lagrangian Approach For CS

$$\min_{x,w} \frac{1}{2}\|H(x)\|_2^2 + \tau\|w\|_1 \qquad \text{such that} \quad \Phi x = w \qquad H(x) = Ax - b$$

Require: $\mu > 0$, tolerance $\tau > 0$, and initial guess $w^0 = 0$, and $\lambda^0 = 0$

while not converged, ie $\|\Phi x^k - w^k\| > \epsilon$ and $\|H(x^k)\|_2^2 > \epsilon$ **do**

Solve L_2 subproblem.

$$x^{k+1} = \min_x \frac{1}{2}\|H(x)\|_2^2 + \frac{1}{2\mu}\|\Phi x - w^k - \mu\lambda^k\|_2^2$$

The derivative of this objective function for the subproblem f_{L2sub} is given by

$$\nabla f_{\text{L2sub}} = A^\top (Ax - b) + \frac{1}{\mu}\Phi^\top \left(\Phi x - w^k - \mu\lambda^k\right)$$

Solve L_1 subproblem.

$$w^{k+1} = \min_w \tau\|w\|_1 + \frac{1}{2\mu}\|\Phi x^{k+1} - w - \mu\lambda^k\|_2^2$$

$$w_i = \text{soft}_{\tau\mu}\left((\Phi x^k)_i - \mu\lambda_i\right)$$

$$\text{soft}_a(b) \equiv \frac{b}{|b|}\max(0, |b| - a)$$

Update Lagrange Multiplier.

$$\lambda^{k+1} = \lambda^k - \frac{1}{\mu}(\Phi x^{k+1} - w^{k+1})$$

end while

Derivatives The L2 subproblem may be solved with a linesearch or trust region approach using finite difference or analytical derivatives. For the general case

$$\min_{\vec{x}\in\mathbb{R}^n} f(\vec{x}) = \min_{\vec{x}\in\mathbb{R}^n} \frac{1}{2}\vec{h}^\top(\vec{x})\vec{h}(\vec{x}) \qquad \vec{h}(\vec{x}) = \begin{bmatrix} A\vec{x} - \vec{b} \\ \frac{1}{\sqrt{\mu}}\left(\Phi\vec{x} - \vec{w} - \mu\vec{\lambda}\right) \end{bmatrix}$$

The gradient of the objective function $f(\vec{x})$ is given as the matrix vector product of the Jacobian transpose times the residual.

$$(\nabla f)_i = \frac{\partial}{\partial x_i}\left(\frac{1}{2}\sum_l^m h_l h_l\right) = \left(\sum_l^m h_l \frac{\partial h_l}{\partial x_i}\right) \qquad \nabla f = J^\top \vec{h} \qquad J_{ij} = \frac{\partial h_i}{\partial x_j} \qquad J = \begin{bmatrix} A \\ \frac{1}{\sqrt{\mu}}\Phi \end{bmatrix}$$

Applying this formula to each term in our L2 subproblem yeilds

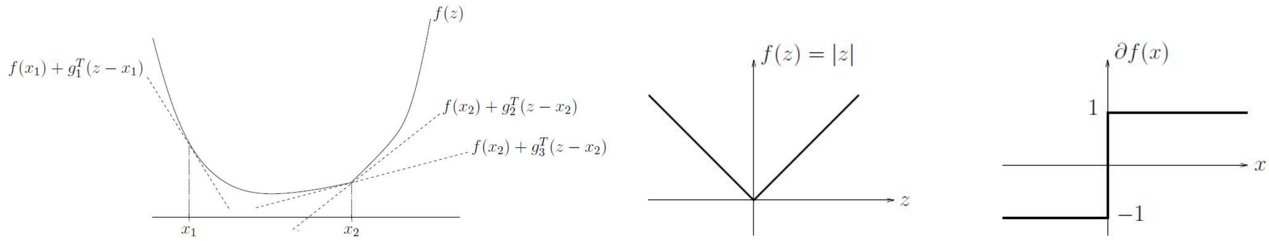
$$\nabla\left(\frac{1}{2}\|H(x)\|_2^2 + \frac{1}{2\mu}\|\Phi x - w^k - \mu\lambda^k\|_2^2\right) = \begin{bmatrix} A \\ \frac{1}{\sqrt{\mu}}\Phi \end{bmatrix}^\top \begin{bmatrix} A\vec{x} - \vec{b} \\ \frac{1}{\sqrt{\mu}}\left(\Phi\vec{x} - \vec{w} - \mu\vec{\lambda}\right) \end{bmatrix} = A^\top(Ax - b) + \frac{1}{\mu}\Phi^\top\left(\Phi x - w^k - \mu\lambda^k\right)$$

Soft Thresholding Operator The L1 subproblem is (perhaps surprisingly) now given by the soft thresholding operator

$$w_i = \text{soft}_{\tau\mu}\left((\Phi x^k)_i - \mu\lambda_i\right) \qquad \text{soft}_a(b) \equiv \frac{b}{|b|}\max(0, |b| - a)$$

To see this, we need to generalize our definition of the derivative for $\|\cdot\|_1$. Consider the subdifferentiable of the L1 problem

$$0 \quad \in \quad \tau\frac{w_i}{|w|_i} - \frac{1}{\mu}\left(((\Phi x^k))_i - w_i - \mu\lambda_i^k\right)$$



(a) (b)

Figure 1: Subdifferential. (a) If f is differentiable, then the gradient is the subgradient. However, the subgradient may exist when the gradient does not exist. In fact there may be several subgradients at this point. (b) Absolute value. Consider $f(z) = |z|$. For $x < 0$ the subgradient is unique: $\partial f(x) = \{-1\}$. Similarly, for $x > 0$ we have $\partial f(x) = \{1\}$. At $x = 0$ the subdifferential is defined by the inequality $|z| > gz$ for all z, which is satisfied if and only if $g \in [-1, 1]$. Therefore we have $\partial f(0) = [-1, 1]$.

Subdifferen

(Definition) Subdifferential We say a vector $g \in \mathbb{R}^n$ is a **subgradient** of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ if

$$f(z) \geq f(x) + g^\top(z - x) \qquad \forall z$$

The **set** of subgradients of f at the point x is called the **subdifferential** at x and is denoted $\partial f(x)$.

$$\partial f(x) \equiv \left\{g : f(z) \geq f(x) + g^\top(z - x) \qquad \forall z\right\}$$

To illustrate the solution to this equation consider

$$0 \in a\ \text{sign}(x) + x - b \quad \Leftrightarrow \quad b \in a\ \text{sign}(x) + x \qquad \text{sign}(x) \equiv \begin{cases} -1 & x < 0 \\ (-1, 1) & x = 0 \\ 1 & x > 0 \end{cases}$$

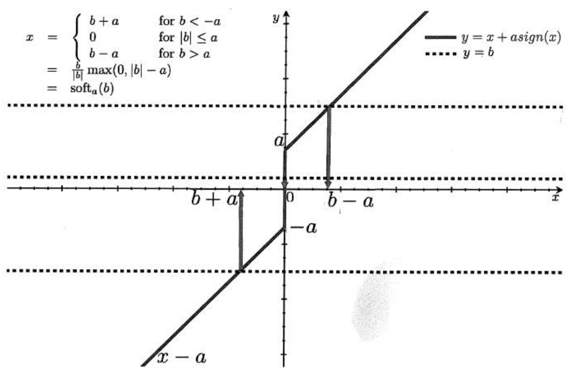


Figure 2: Solution to L_1 subproblem. (Provided by W. Stefan)

As seen in Figure 2, when does $b = x + a\ \text{sign}(x)$?

Case I

$$b > a \quad \Rightarrow \quad b = x + a \quad \Rightarrow \quad x = b - a$$

Case II

$$|b| < a \quad \Rightarrow \quad x = 0$$

Case III

$$b < -a \quad \Rightarrow \quad b = x - a \quad \Rightarrow \quad x = b + a$$

These three cases can be convienently expressed by the so-called soft threshold operator. One may directly verify that:

$$x = \text{soft}_a(b) = \frac{b}{|b|}\max(0, |b| - a) = \begin{cases} b + a & b < -a \\ 0 & |b| \leq -a \\ b - a & b > a \end{cases}$$

The solution to the L1 problem is given by a change of variables.

5 Future Work

From here we plan to investigate applying the FOCUSS algorithm to MRX image formation. This will require improvements on the current assumptions, including a more accurate representation of the sensor geometry. Different initial conditions will need to be investigated including using theoretical values and phantom data. The regularization parameter will need to be optimized. The suitability of other algorithms such as MUSIC and BRAINSTORM will also be considered.

References

[1] Gorodnitsky, et. al. Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm, Electroencephology and clinical Neurophysiology 95 (1995) 231-251.

[Goldstein and Osher, 2009] Goldstein, T. and Osher, S. (2009). The split bregman method for l1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343.