Discretizing the Inverse Problem

Discretizing the Problem

Data eq^{ns} $$\boxed{d_i = (g_i, m) = \int g_i(x)\, m(x)\, dx} \qquad i = 1, N \qquad (1)$$

Also, although we haven't discussed how, we're going to need to evaluate quantities like

$$\phi_s = \int m^2(x)\, dx$$

or $$\phi_x = \int \left(\frac{dm}{dx}\right)^2 dx$$ or higher derivatives

We need to discretize the problem so that it can be solved numerically.

There are numerous options available and the user has a choice. Two routes are

(i) Galerkin

(ii) Quadrature.

Galerkin Methods.

Here we parametrize the model by writing

$$\boxed{m(x) = \sum_{j=1}^{M} \alpha_j \psi_j(x)} \qquad (2)$$

where $\{\psi_j(x)\}$ are expansion basis functions.

eg. (i) Sinusoids
(ii) Box car or pulse basis f^{ns}
(iii) polynomials.

Generally the numerics are helped by using orthonormal (or at least orthogonal) functions, and enough of them. (M>N and also they must be able to capture the model we're looking for)

Substitute (2) into (1)

$$d_i = (g_i, m) = \left(g_i, \sum \alpha_j \psi_j\right) = \sum \alpha_j (g_i, \psi_j)$$

$$d_i = \sum_{j=1}^{M} G_{ij} \alpha_j$$

or $\boxed{d = G\alpha}$ $\qquad G: N \times M$, $\alpha: \mathbb{R}^M$ $\qquad \boxed{G_{ij} = (g_i, \psi_j) = \int g_i(x)\, \psi_j(x)\, dx}$

We can get similar vector matrix expressions for the model norm components.

e.g. $\phi_s = \int m^2(x)\,dx = \int \sum_i \alpha_i \psi_i(x) \sum_j \alpha_j \psi_j(x)\,dx$

$= \sum_i \sum_j \alpha_i \alpha_j (\psi_i, \psi_j)$

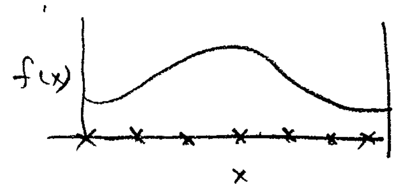
$$\boxed{\phi_s = \alpha^T B \alpha} \qquad \text{where} \quad \boxed{B_{ij} = (\psi_i, \psi_j) = \int \psi_i(x)\psi_j(x)\,dx}$$

Similarly $\quad \phi_x = \alpha^T C \alpha \quad$ where $\quad \boxed{C_{ij} = (\psi_i', \psi_j') = \int \psi_i'(x)\psi_j'(x)\,dx}$

Discretization using Quadrature Integration

Basic quadrature

$$\int_a^b f(x)\,dx = \sum_{k=1}^{M} w_k f(x_k)$$

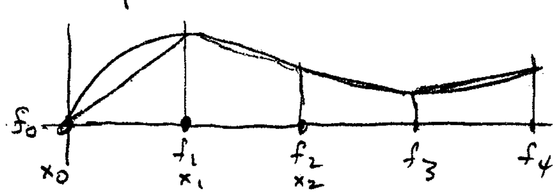


w_k are weights (known)

$f(x_k)$ evaluations of the function at x_k

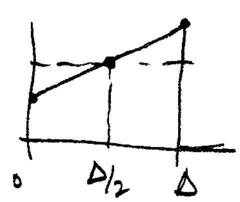
There are many different quadrature formulae that are applicable. (Again the user can choose.)

One that we'll make good use of is the open-ended midpoint rule (effectively the trapezoidal rule for integration.)



Consider $\quad f = ax + b$

$$\int_0^\Delta (ax+b)\,dx = \left.\frac{ax^2}{2} + bx\right|_0^\Delta = \frac{a\Delta^2}{2} + b\Delta = \Delta\left(\frac{a\Delta}{2} + b\right) = \Delta f(\Delta/2)$$

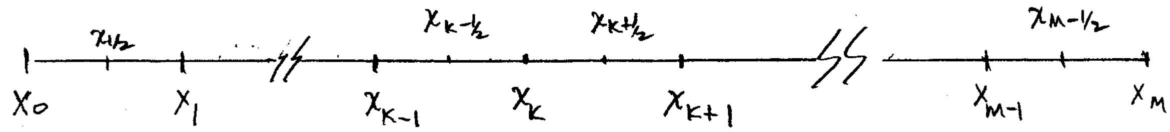


so midpoint rule ≈ trapezoidal rule and is the same if linear interpolation is used

Evaluate $\int_a^b f(x)\,dx$

Introduce a partition $\boxed{a \equiv x_0 < x_1 < \cdots < x_M \equiv b}$

The partition can be uniform or not.



The points $\{x_0, x_1 \cdots x_M\}$ define our <u>primal</u> grid. They often define the boundaries of cells used in our discretization.

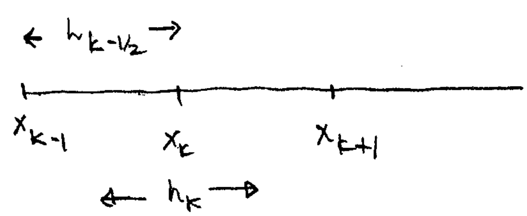
The points $\{x_{1/2}, x_{3/2} \cdots x_{M-1/2}\}$ define our <u>dual</u> grid. They define the centers of cells.

Since we are evaluating integrals using midpoint rules then we need to know width of cells both on our primal grid and also on the dual grid.

Define

$$h_{k-1/2} = x_k - x_{k-1} \qquad 1 \le k \le M$$ length of each cell. (primal grid)

$$h_k = \tfrac{1}{2}(x_{k+1} - x_{k-1}) \equiv \tfrac{1}{2}(h_{k-1/2} + h_{k+1/2})$$ $1 \le k \le M-1$ distance between cell centers



So the subscript locates the length and helps keep track of whether you are looking at the distance between cell centers or the width of the cell.

Data equations:

$$d_j = \int_{x_0}^{x_M} g_j(x)\, m(x)\, dx$$

$$d_j = \sum_{k=1}^{M} g_j(x_{k-1/2})\, h_{k-1/2}\, m(x_{k-1/2})$$

Define a vector of unknowns to be the values of the model at the midpoints of the cells

ie $\vec{m} = \left(m(x_{1/2}), \cdots m(x_{M-1/2}) \right)$

The $\boxed{d = Gm}$

where $\boxed{G_{jk} = g_j(x_{k-1/2})\, h_{k-1/2}}$ $\left(\text{note, this is} \simeq \int_{x_{k-1}}^{x_k} g_j(x)\, dx \text{ as before} \right)$

Model objective f^{n}:

Smallest model component:

$$\phi_m = \int_{x_0}^{x_M} w(x) \left(m(x) - m_{ref}(x) \right)^2 dx$$

$$\boxed{\phi_m = \sum_{k=1}^{M} w(x_{k-1/2}) \left(m(x_{k-1/2}) - m_{ref}(x_{k-1/2}) \right)^2 h_{k-1/2}}$$

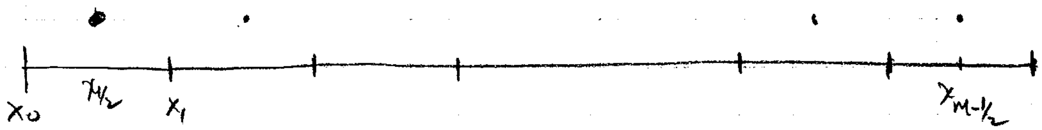
Let $\boxed{W_s = \text{diag}\left(\sqrt{w(x_{k-1/2})\, h_{k-1/2}} \right) \qquad k = 1, \cdots M}$

then $\boxed{\phi_m = \| W_s (m - m_{ref}) \|^2}$ where $m_{ref} \stackrel{def}{=} \left(m_{ref}(x_{k-1/2}) \right)$

Flatteet model

$$\phi_x = \int_{x_0}^{x_M} w(x) \left(\frac{dm}{dx}\right)^2 dx$$

Refer to grid



Our model elements are evaluated at the centers of cells. We can think of forming a derivative only at intervals between these points $[x_{1/2}, x_{M-1/2}]$

So

$$\int_{x_0}^{x_M} \left(\frac{dm}{dx}\right)^2 dx = \int_{x_{1/2}}^{x_{M-1/2}} \left(\frac{dm}{dx}\right)^2 + \underbrace{\int_{x_0}^{x_{1/2}} \left(\frac{dm}{dx}\right)^2 dx + \int_{x_{M-1/2}}^{x_M} \left(\frac{dm}{dx}\right)^2}_{\text{1/2 cells at either end}}$$

we have two options: (1) neglect the contribution from the ends
(2) include them

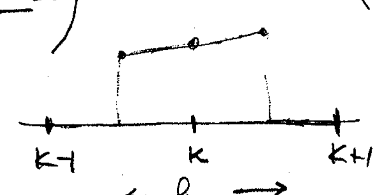
Option 1:

Remember – what is our goal? Want to generate a smooth f^{n}.
If we control the smoothness as evaluated through centers of each cell then this is likely fine.

Generally we don't care about the exact numerical values of ϕ_x (often never look at it) so neglect of contribution to the end points is minimal

Evaluate $$\int_{x_{k-1/2}}^{x_{k+1/2}} \left(\frac{dm}{dx}\right)^2 dx \simeq \left(\frac{m(x_{k+1/2}) - m(x_{k-1/2})}{h_k}\right)^2 h_k \qquad \left(\equiv \left(\frac{dm}{dx}\right)^2_{x_k} h_k\right)$$

is done with the midpoint rule and trapezoidal on the dual grid.



So $$\int_{x_{k-1/2}}^{x_{k+1/2}} w\left(\frac{dm}{dx}\right)^2 dx \simeq \left(m(x_{k+1/2}) - m(x_{k-1/2})\right)^2 \frac{w(x_{k+1/2})}{h_k}$$

The approximation to the full integral becomes

$$\int_{x_0}^{x_M} w\left(\frac{dm}{dx}\right)^2 dx = \sum_{k=1}^{M-1} \frac{w(x_{k+1/2})}{h_k}\left(m(x_{k+1/2}) - m(x_{k-1/2})\right)^2$$

$$\boxed{\phi_x = \|W_x m\|^2}$$

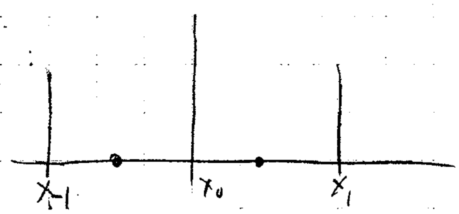
where $$\boxed{W_x = \begin{pmatrix} -\xi_1 & \xi_1 & & & \\ & -\xi_2 & \xi_2 & & \\ & & \ddots & & \\ & & & -\xi_{M-1} & \xi_{M-1} \end{pmatrix}}$$ $$\boxed{\xi_k = \sqrt{\frac{w(x_{k+1/2})}{h_k}}}$$

Now W_x is an $(M-1) \times M$ matrix. It is not full rank.

Option 2: We won't explore this in detail here.

If you were concerned about the contribution from the endpoints then in order to evaluate contributions, you need to have some boundary condition

eg



(i) Setting $$\left.\begin{aligned} m(x_{-1/2}) &= m(1/2) \\ m(x_{M+1/2}) &= m(x_{M-1/2}) \end{aligned}\right\}$$ Newman b.c. $\Rightarrow$ zero contribution. So you can think of what we have done here to be equivalent to that

(ii) Could enforce a Dirichlet b.c., setting $m(x_0)$ or $m(x_M)$. However, this can lead to poor results if you don't have a good estimate

(iii) Also, if you have a bc., you could include $m(x_0)$ as a datum.

(iv) Be careful <u>not</u> to do the following

$$W_x = M{-}1\updownarrow \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}$$ is not invertible

So some people substitute

$$W_x = \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \\ & & & 1 \end{pmatrix}$$ now W_x^{-1} exist

but this is equivalent trying to set $m_M = 0$ (ie a Dirichlet condition)

Remark: What is sometime acceptable if W_x^{-1} is required is to make

$$W_x = \begin{pmatrix} -1 & 1 & & & \\ 0 & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \\ & & & & \epsilon \end{pmatrix}$$ where ϵ is large enough to permit the inverse to be evaluated and small enough so there is no real enforcement of the bc.

Remark: The above work can clearly be altered to incorporate a reference model, so

$$\boxed{\phi_x = \int w(x)\left(\frac{d}{dx}(m - m_{ref})\right)^2 dx = \| W_x (m - m_{ref})\|^2}$$

we have now discretized the inverse problem.

$$\phi_m = \alpha_s \int w_s(x) \left(m(x) - m_{ref}\right)^2 dx + \alpha_x \int w_x(x) \left(\frac{d}{dx}(m - m_{ref})\right)^2 dx$$

$$= \alpha_s \| W_s (m - m_{ref}) \|^2 + \alpha_x \| W_x (m - m_{ref}) \|^2$$

$$= (m - m_{ref})^T \left\{ \alpha_s W_s^T W_s + \alpha_x W_x^T W_x \right\} (m - m_{ref})$$

$$= (m - m_{ref})^T W_m^T W_m (m - m_{ref})$$

So $\boxed{\phi_m = \| W_m (m - m_{ref}) \|^2}$ $\boxed{W_m^T W_m = \alpha_s W_s^T W_s + \alpha_x W_x^T W_x}$

Inverse Problem:

$$\boxed{\begin{array}{ll} \min\limits_{m} & \phi_m = \| W_m (m - m_{ref}) \|^2 \\ \text{subject to} & Gm = d \end{array}}$$

Solution: $\min \; \phi = \| W_m (m - m_{ref}) \|^2 + 2\lambda^T (Gm - d)$

where $\lambda \in \mathbb{R}^N$ is a set of Lagrange multipliers

Solution: $\nabla_m \phi = 0$. and solve

The algebra is simpler if we change variable.

let $x = W_m (m - m_{ref})$

Data eq^n

$$Gm = d$$

$$G(m - m_{ref}) = d - d_{ref} \qquad d_{ref} = G m_{ref}$$

$$\underbrace{G W_m^{-1}}_{A} \underbrace{W_m (m - m_{ref})}_{x} = \underbrace{d - d_{ref}}_{b}$$

Write $A = G W_m^{-1}$ $\quad$ $G : N \times M$ $\quad$ $W_m : M \times M$ $\quad$ W_m^{-1} exists

$b = d - d_{ref}$

Then data eq$\underline{^{ns}}$ are $Ax = b$.

Our minimization problem is

$$\boxed{\begin{aligned} &\text{min.} \quad \phi = \|x\|^2 \\ &\text{subject} \quad Ax = b \end{aligned}}$$

$A : N \times M$

Solve $\quad$ min $\quad \phi = \|x\|^2 + 2\lambda^T(Ax - b)$

$\nabla_x \phi = 0 \qquad 0 = 2x + 2A^T\lambda \quad \Rightarrow \quad \boxed{x = -A^T\lambda} \qquad (1)$

So $\qquad Ax = -AA^T\lambda = b$

but $\qquad AA^T$: $N \times N$ matrix and is invertible

$\Rightarrow \qquad \boxed{\lambda = -(AA^T)^{-1} b} \qquad (2)$

Using (1) yields

$$\boxed{x = A^T (AA^T)^{-1} b} \qquad (3)$$

Since $\quad x = W_m(m - m_{ref})$

$$\boxed{m = W_m^{-1} x + m_{ref}} \qquad (4)$$