MOTIVATION

Numerical problems are everywhere in geometric modeling!

Quick summary!
Mostly for common ground: You may already know this material. First half is important; remainder summarizes interesting recent tools.
Numerical analysis is a huge field.

Patterns, algorithms, & examples common in geometry.
EXAMPLES

How to flatten a mesh?

Object space (3D)  Texture space (2D)

“Mesh Parameterization: Theory and Practice” by Kai Hormann, Bruno Lévy, Alla Sheffer
Credit: Hans-Christian Ebke
EXAMPLES

How to flatten a mesh?

\[
\min_{u_i, v_i \in \mathbb{R}^2} f(u_1, v_1, ..., u_n, v_n)
\]

- Fit a target 2D shape?
- Triangle distortion?
- More sharp corners?
EXAMPLES

How to animate a character?

rest pose  LBS with [Jacobson et al. 2011]  our method

“Linear Subspace Design for Real-Time Shape Deformation” by Yu Wang, Alec Jacobson, Jernej Barbič, Ladislav Kavan
EXAMPLES

How to animate a character?

\[
\min_{x_i, y_i \in \mathbb{R}^2} f(x_1, y_1, \ldots, x_n, y_n)
\]

s.t. \( x_j - x_j' = 0 \)

\( y_j - y_j' = 0 \)

Triangle distortion?  
Some smoothness?  

\( j \in J \)  

Known positions
\[
\min_{x \in \mathbb{R}^n} f(x)
\]
\[
s.t. \ g(x) = 0
\]
\[
h(x) \geq 0
\]
Optimized function  

\( \min_{x \in \mathbb{R}^n} f(x) \)

s.t.  
\( g(x) = 0 \)
\( h(x) \geq 0 \)
\[
\min_{x \in \mathbb{R}^n} f(x)
\]
subject to
\[
g(x) = 0
\]
\[
h(x) \geq 0
\]
\[
\min_{x \in \mathbb{R}^n} f(x)
\]
\[
\text{s.t. } g(x) = 0
\]
\[
\text{h}(x) \geq 0
\]

Inequality constraints
EXAMPLES

\[ Ax = b \iff \min_{x \in \mathbb{R}^n} \| Ax - b \|_2^2 \]
EXAMPLES

\[ Ax = \lambda x \quad \Leftrightarrow \quad \min_{x \in \mathbb{R}^n} \|Ax\|_2^2 \]

\[ \text{s.t.} \quad \|x\| - 1 = 0 \]
ROUGH PLAN

• (intro) Matrices and Eigenvalues
• Linear problems
• Unconstrained optimization
• Equality-constrained optimization
• Variational problems
ROUGH PLAN

• (intro) Matrices and Eigenvalues
• Linear problems
• Unconstrained optimization
• Equality-constrained optimization
• Variational problems
MATRICES

How are those matrices special?

$$\begin{pmatrix} 1 & -2 \\ 5 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & 4 \\ 2 & -4 & 0 \end{pmatrix}$$

$$\begin{pmatrix} \cos(0.4) & -\sin(0.4) & 0 \\ \sin(0.4) & \cos(0.4) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 1 \end{pmatrix}$$
MATRICES

How are those matrices special?

\[
\begin{pmatrix}
1 & -2 \\
5 &
\end{pmatrix}
\]

Diagonal

\[
\begin{pmatrix}
1 & -1 & 0 \\
-1 & 2 & 0 \\
0 & 0 & 5
\end{pmatrix}
\]

Symmetric
\( A = A^T \)

\[
\begin{pmatrix}
0 & -1 & -2 \\
1 & 0 & 4 \\
2 & -4 & 0
\end{pmatrix}
\]

Skew-symmetric
\( A^T = -A \)

\[
\begin{pmatrix}
cos(0.4) & -sin(0.4) & 0 \\
sin(0.4) & cos(0.4) & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

Orthogonal
\( A^{-1} = A^T \)

\[
\begin{pmatrix}
1 & 0 & 3 \\
0 & 1 & 4 \\
& & 1
\end{pmatrix}
\]

Um something?
MATRICES

What is the geometric meaning of those?

\[
\begin{pmatrix}
1 & -2 \\
-2 & 5
\end{pmatrix}
\]

Diagonal

\[
\begin{pmatrix}
\cos(0.4) & -\sin(0.4) & 0 \\
\sin(0.4) & \cos(0.4) & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

Orthogonal

\[
A^{-1} = A^T
\]

Um something?

\[
\begin{pmatrix}
1 & 0 & 3 \\
0 & 1 & 4 \\
0 & 0 & 1
\end{pmatrix}
\]
What is the geometric meaning of those?

\[
\begin{pmatrix}
1 & -2 \\
-2 & 5
\end{pmatrix}
\]

Non-uniform scaling

\[
\begin{pmatrix}
cos(0.4) & -sin(0.4) & 0 \\
sin(0.4) & cos(0.4) & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

Rotation

\[
\begin{pmatrix}
1 & 0 & 3 \\
0 & 1 & 4 \\
0 & 0 & 1
\end{pmatrix}
\]

Translation in 2D*
NO IDEA WHAT THOSE ARE?

https://umontreal.on.worldcat.org/oclc/829434522
Free online at UdeM library
Chapters I.3, I.4
EIGENVALUES AND EIGENVECTORS

\[ Ax = \lambda x \]
EIGENVECTORS/VALUES

Geometric meaning?

$Ax = \lambda x$

Vectors which **only** scale
DIAGONAL MATRICES

• Those are the eigenvalues!
• Scaling along coordinate axes
• $\Theta(n)$ space to store (just $n$ of course)
• $O(n)$ time to invert
SYMMETRIC MATRICES

- All eigenvalues are real
- All eigenvectors are orthogonal

Can be diagonalized by an orthogonal matrix:

$$D = Q^T A Q$$

\[
\begin{pmatrix}
1 & -1 & 0 \\
-1 & 2 & 0 \\
0 & 0 & 5
\end{pmatrix}
\]

Symmetric

$$A = A^T$$
ORTHOGONAL MATRICES

• $AA^T = I$
• Columns have zero dot product
• Doesn’t change length of vectors or angles between them (isometry)
• Determinant is $\pm 1$

\[
\begin{pmatrix}
\cos(0.4) & -\sin(0.4) & 0 \\
\sin(0.4) & \cos(0.4) & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

Orthogonal $A^{-1} = A^T$
MATRICES

• Geometric transformations
• Linear operators on (discrete) functions
MATRIX IS POSITIVE DEFINITE

\[ A \succeq 0 \]

⇔

All eigenvalues are nonnegative

Geometric meaning?
\[
\min_{x \in \mathbb{R}^n} f(x)
\]
\[
\text{s.t. } g(x) = 0
\]
\[
h(x) \geq 0
\]
ROUGH PLAN

• (intro) Matrices and Eigenvalues

• Linear problems

• Unconstrained optimization

• Equality-constrained optimization

• Variational problems
VECTOR SPACES AND LINEAR OPERATORS

\[ L[\vec{x} + \vec{y}] = L[\vec{x}] + L[\vec{y}] \]

\[ L[c\vec{x}] = cL[\vec{x}] \]
EIGENVALUES AND EIGENVECTORS

\[ Ax = \lambda x \]

or

\[ \mathcal{L}[f(x)] = \lambda f(x) \]

Discrete (matrix)

Continuous (operator)
ABSTRACT EXAMPLE

$C^\infty (\mathbb{R})$

$L[f] := \frac{df}{dx}$

$f = e^{3x}$

$L_1 = f' = 3e^{3x}$

$L_2 = f'' = \sin(ax) = a^2 \sin(ax)$

Eigenvectors?
INFINITE DIMENSIONS

\[ A \overset{\rightarrow}{x} \]

matrix vector

\[ \overset{\rightarrow}{x} \mapsto A\overset{\rightarrow}{x} \]

linear operator
LINEAR SYSTEM OF EQUATIONS

\[
\begin{pmatrix}
A
\end{pmatrix}
\begin{pmatrix}
\vec{x}
\end{pmatrix} =
\begin{pmatrix}
\vec{b}
\end{pmatrix}
\]

Simple "inverse problem"
COMMON STRATEGIES

• **Gaussian elimination**
  – $O(n^3)$ time to solve $Ax=b$ or to invert

• **But:** Inversion is unstable and slower!

• **Never ever compute $A^{-1}$ if you can avoid it.**
SIMPLE EXAMPLE

\[ \frac{d^2 f}{dx^2} = g, \quad f(0) = f(1) = 0 \]

Discretization?
On the board
\[
\frac{d^2 f}{dx^2} = g, \quad f(0) = f(1) = 0
\]

\[
\begin{pmatrix}
-2 & 1 & \ \ & \ \ \\
1 & -2 & 1 & \ \ \\
1 & -2 & 1 & \ \ \\
\vdots & \ & \ & \ \\
1 & -2 & 1 & \ \\
1 & -2 & \ \ & \ \\
\end{pmatrix}
\begin{pmatrix}
f_1 \\
f_2 \\
\vdots \\
f_n \\
\end{pmatrix}
= 
\begin{pmatrix}
g_1 \\
g_2 \\
\vdots \\
g_n \\
\end{pmatrix}
\]
STRUCTURE?

$$\begin{pmatrix} -2 & 1 & \cdots & 1 \\ 1 & -2 & 1 & \cdots \\ 1 & -2 & 1 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 1 & -2 & 1 & \cdots \\ 1 & -2 & 1 & -2 \end{pmatrix}$$
LINEAR SOLVER
CONSIDERATIONS

• Never construct $A^{-1}$ explicitly
  (if you can avoid it)

• Added structure helps
  Sparsity, symmetry, positive definiteness, bandedness

$$\text{inv}(A) \times b \ll (A^*A) \backslash (A^*b) \ll A \backslash b$$
LINEAR SYSTEMS: SOLVERS

• **Direct** *(explicit matrix)*
  – **Dense**: Gaussian elimination/LU, QR for least-squares
  – **Sparse**: Reordering (SuiteSparse, Eigen)

• **Iterative** *(apply matrix repeatedly)*
  – **Positive definite**: Conjugate gradients
  – **Symmetric**: MINRES, GMRES
  – **Generic**: LSQR
GENERIC ADVICE

Generic tools are often not too effective!
 GENERIC ADVICE

Try the simplest solver first.
VERY COMMON: SPARSITY

Induced by the connectivity of the triangle mesh.

Iteration of CG has local effect ⇒ Precondition!
• No need to implement a linear solver

• If a matrix is sparse, your code should store it as a sparse matrix!

https://eigen.tuxfamily.org/dox/group__TutorialSparse.html
ROUGH PLAN

• *(intro)* Matrices and Eigenvalues
• Linear problems
• Unconstrained optimization
• Equality-constrained optimization
• Variational problems
UNCONSTRAINED OPTIMIZATION

\[ \min_{x \in \mathbb{R}^n} f(x) \]

Trivial when \( f(x) \) is linear

Easy when \( f(x) \) is quadratic

Hard in case of generic non-linear.
UNCONSTRAINED OPTIMIZATION

\[
\min_{x \in \mathbb{R}^n} f(x)
\]
NOTIONS FROM CALCULUS

\[ f(x, y) = 3x^2 - y \]
\[ \nabla f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (6x, -1) \]

Geometric meaning?
NOTIONS FROM CALCULUS

\[ f : \mathbb{R}^n \rightarrow \mathbb{R} \]

\[ \nabla f = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right) \]

https://en.wikipedia.org/?title=Gradient

Gradient
NOTIONS FROM CALCULUS

\[ f: \mathbb{R}^n \rightarrow \mathbb{R} \]

\[ \nabla f = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right) \]

https://en.wikipedia.org/?title=Gradient
NOTIONS FROM CALCULUS

\[ f : \mathbb{R}^n \rightarrow \mathbb{R}^m \]

\[ \rightarrow (Df)_{ij} = \frac{\partial f_i}{\partial x_j} \]

NOTIONS FROM CALCULUS

\[ f : \mathbb{R}^n \rightarrow \mathbb{R} \rightarrow H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} \]

\[ f(x) \approx f(x_0) + \nabla f(x_0)^\top (x - x_0) + (x - x_0)^\top H f(x_0)(x - x_0) \]

http://math.etsu.edu/multicalc/prealpha/Chap2/Chap2-5/10-3a-t3.gif

Hessian
OPTIMIZATION TO ROOT-FINDING

$$\nabla f(x) = 0$$

(unconstrained)

Critical point

Saddle point

Local min

Local max

$$f(x)$$

$$x$$

Critical point
CONVEX FUNCTIONS

\[ f''(x) > 0 \]

https://en.wikipedia.org/wiki/Convex_function
CONVEX FUNCTIONS

\[ H(x) \geq 0 \]

https://en.wikipedia.org/wiki/Convex_function
SPECIAL CASE: LEAST-SQUARES

\[
\min_x \frac{1}{2} \| Ax - b \|_2^2
\]

\[
\rightarrow \min_x \frac{1}{2} x^\top A^\top A x - b^\top A x + \| b \|_2^2
\]

\[
\implies A^\top A x = A^\top b
\]

**Normal equations**

*(better solvers for this case!)*
USEFUL DOCUMENT

The Matrix Cookbook
Petersen and Pedersen

UNCONSTRAINED OPTIMIZATION

\[
\min_x f(x)
\]

Unstructured.
BASIC ALGORITHMS

Gradient descent:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

Multiple optima!

Line search
BASIC ALGORITHMS

\[ \lambda_0 = 0, \lambda_s = \frac{1}{2}(1 + \sqrt{1 + 4\lambda_{s-1}^2}), \gamma_s = \frac{1 - \lambda^2}{\lambda_{s+1}} \]

\[ y_{s+1} = x_s - \frac{1}{\beta} \nabla f(x_s) \]

\[ x_{s+1} = (1 - \gamma_s)y_{s+1} + \gamma_s y_s \]

Quadratic convergence on convex problems!
(Nesterov 1983)

A very cool intro: https://distill.pub/2017/momentum/

Accelerated gradient descent
Newton’s Method

\[ x_{k+1} = x_k - \left[ H f(x_k) \right]^{-1} \nabla f(x_k) \]

**Line search for stability**
BASIC ALGORITHMS

\[ x_{k+1} = x_k - M_k^{-1} \nabla f(x_k) \]

- (Often \textit{sparse}) approximation from previous samples and gradients
- Inverse in \textit{closed form}!

Quasi-Newton: BFGS and friends
EXAMPLE: SHAPE INTERPOLATION

\[ x(t) = (1 - t) x_0 + t x, \quad t = 0 \quad t = 1 \]

Figure 5: Interpolation and extrapolation of the yellow example poses. The blending weights are 0, 0.35, 0.65, 1.0, and 1.25.

Figure 6: Interpolation of an adaptively meshed and strongly twisted helix with blending weights 0, 0.25, 0.5, 0.75, 1.0.

Fröhlich and Botsch. “Example-Driven Deformations Based on Discrete Shells.” CGF 2011.
INTERPOLATION PIPELINE

Roughly:

1. **Linearly interpolate** edge lengths and dihedral angles.

   \[
   \ell_e^* = (1 - t)\ell_e^0 + t\ell_e^1 \\
   \theta_e^* = (1 - t)\theta_e^0 + t\theta_e^1
   \]

2. **Nonlinear optimization** for vertex positions.

   \[
   \min_{x_1, \ldots, x_m} \lambda \sum_e w_e (\ell_e(x) - \ell_e^*)^2 \\
   + \mu \sum_e w_b (\theta_e(x) - \theta_e^*)^2
   \]

**Sum of squares:** Gauss–Newton
SOFTWARE

- Matlab: fminunc or minfunc
- C++: libLBFGS, dlib, others

Typically provide functions for **function and gradient** (and optionally, Hessian).

Try several!
SOME TRICKS

Lots of small elements: \( \|x\|_2^2 = \sum_i x_i^2 \)

Lots of zeros: \( \|x\|_1 = \sum_i |x_i| \)

Uniform norm: \( \|x\|_{\infty} = \max_i |x_i| \)

Low rank: \( \|X\|_* = \sum \sigma_i \)

Mostly zero columns: \( \|X\|_{2,1} = \sum_j \sqrt{\sum_i x_{ij}^2} \)

Smooth: \( \int \|\nabla f\|_2^2 \)

Piecewise constant: \( \int \|\nabla f\|_2 \)

???: Early stopping

Regularization
SOME TRICKS

Multiscale/graduated optimization
ROUGH PLAN

• (intro) Matrices and Eigenvalues
• Linear problems
• Unconstrained optimization
• Equality-constrained optimization
• Variational problems
LAGRANGE MULTIPLIERS: IDEA

\[
\min_x f(x) \\
\text{s.t. } g(x) = 0
\]
LAGRANGE MULTIPLIERS: IDEA

\[ \min_x \ f(x) \]
\[ \text{s.t.} \quad g(x) = 0 \]

- Decrease \( f \): \(-\nabla f\)
- Violate constraint: \( \pm \nabla g \)
LAGRANGE MULTIPLIERS: IDEA

\[
\min_x f(x) \\
\text{s.t. } g(x) = 0
\]

Want:

\[\nabla f \parallel \nabla g\]

\[\implies \nabla f = \lambda \nabla g\]
USE OF LAGRANGE MULTIPLIERS

Turns constrained optimization into unconstrained root-finding.

\[ \nabla f(x) = \lambda \nabla g(x) \]
\[ g(x) = 0 \]
\[ \min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad g(x) = 0 \]
QUADRATIC WITH LINEAR EQUALITY

\[
\begin{align*}
\min_x & \quad \frac{1}{2} x^T A x - b^T x + c \\
\text{s.t.} & \quad M x = v
\end{align*}
\]
(assume A is symmetric and positive definite)

\[
\nabla f = \lambda \nabla g
\]

\[
\nabla f = A x - b = M^T \lambda
\]

\[
g = 0 \quad \Rightarrow \quad \left( \begin{array}{cc} A & M^T \\ M & 0 \end{array} \right) \left( \begin{array}{c} x \\ \lambda \end{array} \right) = \left( \begin{array}{c} b \\ v \end{array} \right)
\]
QUADRATIC WITH LINEAR EQUALITY

\[
\begin{align*}
\min_x & \quad \frac{1}{2} x^\top A x - b^\top x + c \\
\text{s.t.} & \quad M x = v
\end{align*}
\]

(assume \( A \) is symmetric and positive definite)

\[
\begin{pmatrix}
A & M^\top \\
M & 0
\end{pmatrix}
\begin{pmatrix}
x \\
\lambda
\end{pmatrix}
= 
\begin{pmatrix}
b \\
v
\end{pmatrix}
\]
\[
\min_{x \in \mathbb{R}^n} f(x)
\text{ s.t. } g(x) = 0
\]
MANY OPTIONS

• Reparameterization
  Eliminate constraints to reduce to unconstrained case

• Newton’s method
  Approximation: quadratic function with linear constraint

• Penalty method
  Augment objective with barrier term, e.g. $f(x) + \rho |g(x)|$
EXAMPLE: SYMMETRIC EIGENVECTORS

\[ f(x) = x^\top Ax \implies \nabla f(x) = 2Ax \]

\[ g(x) = \|x\|^2_2 \implies \nabla g(x) = 2x \]

\[ \implies Ax = \lambda x \]
EXAMPLE: MESH EMBEDDING

G. Peyré, mesh processing course slides
LINEAR SOLVE FOR EMBEDDING

\[ x_i \in \mathbb{R}^2 \]

\[
\min_{x_1, \ldots, x_{|V|}} \sum_{(i,j) \in E} w_{ij} \| x_i - x_j \|^2_2 \\
\text{s.t. } x_v \text{ fixed } \forall v \in V_0
\]

- \( w_{ij} \equiv 1 \): Tutte embedding
- \( w_{ij} \text{ from mesh} \): Harmonic embedding

Assumption: \( w \) symmetric.
What if $V_0 = \emptyset$?
NONTRIVIALITY CONSTRAINT

\[
\begin{array}{l}
\min_x \|Ax\|_2 \\
\text{s.t.} \|x\|_2 = 1
\end{array}
\quad \mapsto \quad A^\top Ax = \lambda x
\]

Prevents trivial solution \( x \equiv 0 \).

Extract the smallest eigenvalue.

\[
\begin{align*}
\min_u & \quad u^\top L_C u \\
\text{s.t.} & \quad u^\top B e = 0 \quad \text{(Easy fix)} \\
& \quad u^\top B u = 1
\end{align*}
\]
BASIC IDEA OF EIGENALGORITHMS

\[ A\vec{v} = c_1 A\vec{x}_1 + \cdots + c_n A\vec{x}_n \]
\[ = c_1 \lambda_1 \vec{x}_1 + \cdots + c_n \lambda_n \vec{x}_n \text{ since } A\vec{x}_i = \lambda_i \vec{x}_i \]
\[ = \lambda_1 \left( c_1 \vec{x}_1 + \frac{\lambda_2}{\lambda_1} c_2 \vec{x}_2 + \cdots + \frac{\lambda_n}{\lambda_1} c_n \vec{x}_n \right) \]
\[ A^2 \vec{v} = \lambda_1^2 \left( c_1 \vec{x}_1 + \left( \frac{\lambda_2}{\lambda_1} \right)^2 c_2 \vec{x}_2 + \cdots + \left( \frac{\lambda_n}{\lambda_1} \right)^2 c_n \vec{x}_n \right) \]
\[ \vdots \]
\[ A^k \vec{v} = \lambda_1^k \left( c_1 \vec{x}_1 + \left( \frac{\lambda_2}{\lambda_1} \right)^k c_2 \vec{x}_2 + \cdots + \left( \frac{\lambda_n}{\lambda_1} \right)^k c_n \vec{x}_n \right) . \]
TRUST REGION METHODS

Example: Levenberg-Marquardt

\[
\begin{align*}
\min_{\delta x} & \quad \frac{1}{2} \delta x^\top H \delta x + w^\top x \\
\text{s.t.} & \quad \|\delta x\|_2^2 \leq \Delta \\
\downarrow & \\
(H + \lambda I)\delta x &= -w
\end{align*}
\]

Fix (or adjust) damping parameter $\lambda > 0$. 

$\Delta$
EXAMPLE: POLYCUDE MAPS


Align with coordinate axes

\[
\min_X \sum_{b_i} A(b_i; X) \| n(b_i; X) \|_1
\]

\[
\sum_{b_i} A(b_i; X) = \sum_{b_i} A(b_i; X_0)
\]

Preserve area

Note: Final method includes more terms!
Aside:
Convex Optimization Tools

Sometimes work for non-convex problems...

Try lightweight options
ITERATIVELY REWEIGHTED LEAST SQUARES

\[
\min_x \sum_i \phi(x^\top a_i + b_i) \iff \begin{cases} 
\min_{x,y_i} \sum_i y_i (x^\top a_i + b_i)^2 \\
\text{s.t. } y_i = \phi(x^\top a_i + b_i)(x^\top a_i + b_i)^{-2} 
\end{cases}
\]

“Geometric median”

\[
\min_x \sum_i \|x - p_i\|_2 \implies \begin{cases} 
x \leftarrow \min_x \sum_i y_i \|x - p_i\|_2^2 \\
y_i \leftarrow \|x - p_i\|_2^{-1}
\end{cases}
\]

Repeatedly solve linear systems
ALTERNATING PROJECTION

\[
\min_p d(p, p_0) \\
\text{s.t. } p \in C_1 \cap C_2 \cap \cdots \cap C_k
\]

\text{\textcolor{red}{d can be a Bregman divergence}}
AUGMENTED LAGRANGIANS

\[
\begin{align*}
\min_x & \quad f(x) \\
\text{s.t.} & \quad g(x) = 0 \\
\downarrow & \\
\min_x & \quad f(x) + \frac{\rho}{2} \|g(x)\|_2^2 \\
\text{s.t.} & \quad g(x) = 0
\end{align*}
\]

Does nothing when constraint is satisfied

Add constraint to objective
ALTERNATING DIRECTION METHOD OF MULTIPLIERS (ADMM)

\[ \min_{x, z} \ f(x) + g(z) \]

\[ \text{s.t.} \quad Ax + Bz = c \]

\[ \Lambda_\rho(x, z; \lambda) = f(x) + g(z) + \lambda^\top (Ax + Bz - c) + \frac{\rho}{2} \|Ax + Bz - c\|_2^2 \]

\[ x \leftarrow \arg \min_x \Lambda_\rho(x, z, \lambda) \]

\[ z \leftarrow \arg \min_z \Lambda_\rho(x, z, \lambda) \]

\[ \lambda \leftarrow \lambda + \rho(Ax + Bz - c) \]
FRANK-WOLFE

To minimize $f(x)$ s.t. $x \in \mathcal{D}$:

$$s_k \leftarrow \left\{ \begin{array}{l} \text{arg} \min_s \ s^\top \nabla f(x_k) \\ \text{s.t. } s \in \mathcal{D} \end{array} \right\}$$

$$\gamma \leftarrow \frac{2}{k + 2}$$

$$x_{k+1} \leftarrow x_k + \gamma(s_k - x_k)$$

https://en.wikipedia.org/wiki/Frank%E2%80%93Wolfe_algorithm
ROUGH PLAN

• *(intro)* Matrices and Eigenvalues
• Linear problems
• Unconstrained optimization
• Equality-constrained optimization
• Variational problems
VARIATIONAL CALCULUS: BIG IDEA

Sometimes your unknowns are not numbers!

Can we use calculus to optimize anyway?
ON THE BOARD

\[
\min \int_{\Omega} \| \vec{v}(x) - \nabla f(x) \|_2^2 \, d\vec{x}
\]

\[
\min \int_{\Omega} f(x)^2 \, d\vec{x} = 1 \int_{\Omega} \| \nabla f(x) \|_2^2 \, d\vec{x}
\]
GÂTEAUX DERIVATIVE

$$dF[u; \psi] := \frac{d}{dh} F[u + h\psi] \big|_{h=0}$$

Vanishes for all $\psi$ at a critical point!

Analog of derivative at $u$ in $\psi$ direction
“I got a great reputation for doing integrals, only because my box of tools was different from everybody else's, and they had tried all their tools on it before giving the problem to me.”

Richard P. Feynman