Applications

Lijun Zhang
zlj@nju.edu.cn
http://cs.nju.edu.cn/zlj

Outline

- Norm Approximation
  - Basic Norm Approximation
  - Approximation with Constraints
- Least-norm Problems
- Regularized Approximation
- Projection
  - Projection on a Set
  - Projection on a Convex Set

Basic Norm Approximation

- Norm Approximation Problem

$$\min \quad \|Ax - b\|$$

  - $A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$ are problem data
  - $x \in \mathbf{R}^n$ is the variable
  - $\|\cdot\|$ is a norm on $\mathbf{R}^n$
  - Approximation solution of $Ax \approx b$, in $\|\cdot\|$
- Residual

$$r = Ax - b$$

- A Convex Problem
  - $b \in \mathcal{R}(A)$, the optimal value is 0
  - $b \notin \mathcal{R}(A)$, more interesting ($m > n$)

Basic Norm Approximation

Approximation Interpretation

$$Ax = x_1 a_1 + \cdots + x_n a_n$$

- $a_1, \ldots, a_n \in \mathbf{R}^m$ are the columns of A
- Approximate the vector b by a linear combination

- Regression problem
  - ✓ $a_1, \ldots, a_n$ are regressors
  - ✓ $x_1 a_1 + \cdots + x_n a_n$ is the regression of b

Basic Norm Approximation

Estimation Interpretation

- Consider a linear measurement model

$$y = Ax + v$$

- $y \in \mathbf{R}^m$ is a vector measurement
- $x \in \mathbf{R}^n$ is a vector of parameters to be estimated
- $v \in \mathbf{R}^m$ is some measurement error that is unknown, but presumed to be small
- Assume smaller values of v are more plausible

$$\hat{x} = \operatorname{argmin}_z \|Az - y\|$$

Basic Norm Approximation

Geometric Interpretation

- Consider the subspace $\mathcal{A} = \mathcal{R}(A) \subseteq \mathbf{R}^m$, and a point $b \in \mathbf{R}^m$
- A projection of the point b onto the subspace $\mathcal{A}$, in the norm $\|\cdot\|$

$$\begin{array}{ll} \min & \|u - b\| \\ \text{s.t.} & u \in \mathcal{A} \end{array}$$

- Parametrize an arbitrary element of $\mathcal{R}(A)$ as $u = Ax$, we see that norm approximation is equivalent to projection

Basic Norm Approximation

Least-Squares Approximation

$$\min \quad \|Ax-b\|_2^2 = r_1^2 + r_2^2 + \cdots + r_m^2$$

- The minimization of a convex quadratic function

$$f(x) = x^\top A^\top A x - 2b^\top A x + b^\top b$$

- A point x minimizes f if and only if

$$\nabla f(x) = 2A^\top A x - 2A^\top b = 0$$

- Normal equations

$$A^\top A x = A^\top b$$

Basic Norm Approximation

- Chebyshev or Minimax Approximation

$$\min \quad \|Ax-b\|_\infty = \max\{|r_1|, \dots, |r_m|\}$$

  - Be cast as an LP

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & -t\mathbf{1} \preccurlyeq Ax-b \preccurlyeq t\mathbf{1} \end{aligned}$$

  with variables $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$

- Sum of Absolute Residuals Approximation

$$\min \quad \|Ax-b\|_1 = |r_1| + \cdots + |r_m|$$

  - Be cast as an LP

$$\begin{aligned} \min \quad & \mathbf{1}^\top t \\ \text{s.t.} \quad & -t \preccurlyeq Ax-b \preccurlyeq t \end{aligned}$$

  with variables $x \in \mathbf{R}^n$ and $t \in \mathbf{R}^m$

Outline

- Norm Approximation
  - Basic Norm Approximation
  - Approximation with Constraints
- Least-norm Problems
- Regularized Approximation
- Projection
  - Projection on a Set
  - Projection on a Convex Set

Approximation with Constraints

- Add Constraints to

$$\min \quad \|Ax - b\|$$

  - Rule out certain unacceptable approximations of the vector b
  - Ensure that the approximator Ax satisfies certain properties
  - Prior knowledge of the vector x to be estimated
  - Prior knowledge of the estimation error v
  - Determine the projection of a point b on a set more complicated than a subspace

Approximation with Constraints

- Nonnegativity Constraints on Variables

$$\begin{array}{ll} \min & \|Ax - b\| \\ \text{s.t.} & x \succcurlyeq 0 \end{array}$$

  - Estimate a vector x of parameters known to be nonnegative
  - Determine the projection of a vector b onto the cone generated by the columns of A
  - Approximate b using a nonnegative linear combination of the columns of A

Approximation with Constraints

Variable Bounds

$$\begin{aligned} \min \quad & \|Ax - b\| \\ \text{s.t.} \quad & l \leqslant x \leqslant u \end{aligned}$$

- Prior knowledge of intervals in which each variable lies

- Determine the projection of a vector b onto the image of a box under the linear mapping induced by A

Approximation with Constraints

Probability Distribution

$$\begin{aligned} \min \quad & \|Ax - b\| \\ \text{s.t.} \quad & x \succcurlyeq 0, 1^\top x = 1 \end{aligned}$$

- Estimation of proportions or relative frequencies
- Approximate b by a convex combination of the columns of A

Norm Ball Constraint

$$\begin{aligned} \min \quad & \|Ax - b\| \\ \text{s.t.} \quad & \|x - x_0\| \leq d \end{aligned}$$

- x_0 is prior guess of what the parameter x is, and d is the maximum plausible deviation

Outline

- Norm Approximation
  - Basic Norm Approximation
  - Approximation with Constraints
- Least-norm Problems
- Regularized Approximation
- Projection
  - Projection on a Set
  - Projection on a Convex Set

Least-norm Problems

□ Basic least-norm Problem

$$\begin{aligned} \min \quad & \|x\| \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

- $A \in \mathbf{R}^{m \times n}, b \in \mathbf{R}^m$
- $x \in \mathbf{R}^n$, $\|\cdot\|$ is a norm on $\mathbf{R}^n$
- The solution is called a least-norm solution of $Ax = b$
- A convex optimization problem
- Interesting when $m < n$
  - ✓ When the equation is underdetermined

Least-norm Problems

- Reformulation as Norm Approximation Problem
  - Let x_0 be any solution of $Ax = b$
  - Let $Z \in \mathbf{R}^{n \times k}$ be a matrix whose columns are a basis for the nullspace of A.

$$\{x | Ax = b\} = \{x_0 + Zu | u \in \mathbf{R}^k\}$$

  - The least-norm problem can be expressed as

$$\min \quad \|x_0 + Zu\|$$

Least-norm Problems

Estimation Interpretation

- We have $m < n$ perfect linear measurement, given by $Ax = b$
- Our measurements do not completely determine x

- Suppose our prior information, is that x is more likely to be small than large
- Choose the parameter vector x which is smallest among all parameter vectors that are consistent with the measurements

Least-norm Problems

- Geometric Interpretation
  - The feasible set $\{x|Ax = b\}$ is affine
  - The objective is the distance between x and the point 0
  - Find the point in the affine set with minimum distance to 0
  - Determine the projection of the point 0 on the affine set $\{x|Ax = b\}$

Least-norm Problems

- Least-squares Solution of Linear Equations

$$\begin{aligned} \min \quad & \|x\|_2^2 \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

  - The optimality conditions

$$2x^* + A^\top v^* = 0 \quad Ax^* = b$$

    - ✓ v is the dual variable

  - The Solution

$$x^* = -\frac{1}{2}A^\top v^* \Rightarrow -\frac{1}{2}AA^\top v^* = b$$

$$\Rightarrow \quad v^* = -2(AA^\top)^{-1}b, x^* = A^\top(AA^\top)^{-1}b$$

Least-norm Problems

- Sparse Solutions via Least ℓ_1-norm

$$\begin{aligned} \min \quad & \|x\|_1 \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

  - Tend to produce a solution x with a large number of components equal to 0
  - Tend to produce sparse solutions of $Ax = b$, often with m nonzero components

Outline

- Norm Approximation
  - Basic Norm Approximation
  - Approximation with Constraints
- Least-norm Problems
- Regularized Approximation
- Projection
  - Projection on a Set
  - Projection on a Convex Set

Bi-criterion Formulation

- A (convex) Vector Optimization Problem with Two Objectives

$$\min(\text{w.r.t.}\ \mathbf{R}_+^2) \quad (\|Ax-b\|, \|x\|)$$

  - Find a vector x that is small
  - Make the residual $Ax-b$ small
  - Optimal trade-off between the two objectives
    - ✓ The minimum value of $\|x\|$ is 0 and the residual norm is $\|b\|$
    - ✓ Let C denote the set of minimizers of $\|Ax-b\|$, and then any minimum norm point in C is Pareto optimal

Regularization

- Weighted Sum of the Objectives

$$\min \quad \|Ax - b\| + \gamma\|x\|$$

  - $\gamma > 0$ is a problem parameter
  - A common scalarization method used to solve the bi-criterion problem
  - As γ varies over $(0, \infty)$, the solution traces out the optimal trade-off curve

- Weighted Sum of Squared Norms

$$\min \quad \|Ax - b\|^2 + \gamma\|x\|^2$$

Regularization

☐ Tikhonov Regularization

$$\min \quad \|Ax - b\|_2^2 + \delta\|x\|_2^2 = x^\top(A^\top A + \delta I)x - 2b^\top Ax + b^\top b$$

- Analytical solution

$$x = (A^\top A + \delta I)^{-1}A^\top b$$

- Since $A^\top A + \delta I \succ 0$ for any $\delta \succ 0$, the Tikhonov regularized least-squares solution requires no rank assumptions on the matrix A

Outline

- Norm Approximation
  - Basic Norm Approximation
  - Approximation with Constraints
- Least-norm Problems
- Regularized Approximation
- Projection
  - Projection on a Set
  - Projection on a Convex Set

Projection on a Set

- The distance of a point $x_0 \in \mathbf{R}^n$ to a closed set $C \subseteq \mathbf{R}^n$, in the norm $\|\cdot\|$

$$\mathrm{dist}(x_0, C) = \inf\{\|x_0 - x\| \mid x \in C\}$$

  - The infimum is always achieved
- Projection of x_0 on C
  - Any point $z \in C$ which is closest to x_0

$$\|z - x_0\| = \mathrm{dist}(x_0, C)$$

  - Can be more than one projection of x_0 on C
  - If C is closed and convex, and the norm is strictly convex, there is exactly one

Projection on a Set

- The distance of a point $x_0 \in \mathbf{R}^n$ to a closed set $C \subseteq \mathbf{R}^n$, in the norm $\|\cdot\|$

$$\mathrm{dist}(x_0, C) = \inf\{\|x_0 - x\| \mid x \in C\}$$

  - The infimum is always achieved

- $P_C: \mathbf{R}^n \longrightarrow \mathbf{R}^n$ to denote the projection of x_0 on C

$$P_C(x_0) \in C, \|x_0 - P_C(x_0)\| = \mathrm{dist}(x_0, C)$$

$$P_C(x_0) = \mathrm{argmin}\{\|x - x_0\| \mid x \in C\}$$

  - We refer to P_C as projection on C

Example

- Projection on the Unit Square in $\mathbf{R}^2$
  - Consider the boundary of the unit square in $\mathbf{R}^2$, i.e., $C = \{x \in \mathbf{R}^2 \mid \|x\|_\infty = 1\}$, take $x_0 = 0$
  - In the ℓ_1-norm, the four points $(1,0)$, $(0,-1)$, $(-1,0)$, and $(0,1)$ are closest to $x_0 = 0$, with distance 1, so we have $\mathrm{dist}(x_0, C) = 1$ in the ℓ_1-norm
  - In the ℓ_∞-norm, all points in C lie at a distance 1 from x_0, and $\mathrm{dist}(x_0, C) = 1$

Example

Projection onto Rank-k Matrices

- The set of $m \times n$ matrices with rank less than or equal to k

$$C = \{X \in \mathbf{R}^{m\times n} | \operatorname{rank} X \leq k\}$$

  with $k \leq \min\{m, n\}$

- The Projection of $X_0 \in \mathbf{R}^{m\times n}$ on C in $\|\cdot\|_2$
  - ✓ SVD of X_0

$$X_0 = \sum_{i=1}^{r} \sigma_i u_i v_i^\top$$

$$P_C(X_0) = \sum_{i=1}^{\min\{k,r\}} \sigma_i u_i v_i^\top$$

Outline

- Norm Approximation
  - Basic Norm Approximation
  - Approximation with Constraints
- Least-norm Problems
- Regularized Approximation
- Projection
  - Projection on a Set
  - Projection on a Convex Set

Projection on a Convex Set

- C is Convex
  - Represent C by a set of linear equalities and convex inequalities

$$Ax = b, \qquad f_i(x) \le 0, i = 1, \dots, m$$

- Projection of x_0 on C

$$\begin{aligned} \min \quad & \|x - x_0\| \\ \text{s.t.} \quad & f_i(x) \le 0, i = 1, \dots, m \\ & Ax = b \end{aligned}$$

  - A convex optimization problem
  - Feasible if and only if C is nonempty

Example

Euclidean Projection on a Polyhedron

- Projection of x_0 on $C = \{x | Ax \preccurlyeq b\}$

$$\begin{array}{ll} \min & \|x - x_0\|_2^2 \\ \text{s.t.} & Ax \preccurlyeq b \end{array}$$

- Projection of x_0 on $C = \{x | a^\top x = b\}$

$$P_C(x_0) = x_0 + \frac{(b - a^\top x_0)a}{\|a\|_2^2}$$

- Projection of x_0 on $C = \{x | a^\top x \leq b\}$

$$P_C(x_0) = \begin{cases} x_0 + \dfrac{(b - a^\top x_0)a}{\|a\|_2^2}, & a^\top x_0 > b \\ x_0, & a^\top x_0 \leq b \end{cases}$$

Example

Euclidean Projection on a Polyhedron

- Projection of x_0 on $C = \{x | l \preccurlyeq x \preccurlyeq u\}$

$$P_C(x_0)_k = \begin{cases} l_k, & x_{0k} \leq l_k \\ x_{0k}, & l_k \leq x_{0k} \leq u_k \\ u_k, & u_k \leq x_{0k} \end{cases}$$

Property of Euclidean Projection

- C is Convex

$$\|P_C(x) - P_C(y)\|_2 \leq \|x - y\|_2$$

for all x, y

Example

- $K = \mathbf{R}_+^n$

$$P_K(x_0)_k = \max\{x_{0k}, 0\}$$

  - Replace each negative component with 0

- $K = \mathbf{S}_+^n$ and $\|\cdot\|_F$

$$P_K(X_0) = \sum_{i=1}^{n} \max\{0, \lambda_i\} v_i v_i^\top$$

  - The eigendecomposition of X_0 is $X_0 = \sum_{i=1}^{n} \lambda_i v_i v_i^\top$
  - Drop terms associated with negative eigenvalues

Outline

- Norm Approximation
  - Basic Norm Approximation
  - Approximation with Constraints
- Least-norm Problems
- Regularized Approximation
- Projection
  - Projection on a Set
  - Projection on a Convex Set

Homework 3

- http://www.lamda.nju.edu.cn/qiuzh/optfall2021gra.html
- Due: Dec 21, at 11:59 PM
  - 最后一次上课前一天