Lecture 12: Global Optimization

1 Projective Least Squares

In Lecture 9 we studied local optimization methods for multiple view geometry problems. Under the assumption of Gaussian image noise we optimized the maximum likelihood function

$$\sum_{i,j}(r_{ij}(v))^2 = \sum_{ij}\left\|\left(x_{ij}^1 - \frac{R_i^1 X_j + t_i^1}{R_i^3 X_j + t_i^3}, x_{ij}^2 - \frac{R_i^2 X_j + t_i^2}{R_i^3 X_j + t_i^3}\right)\right\|^2. \quad (1)$$

In general this problem is difficult to solve since the involved terms are quotients of quadratic functions. In this lecture we will study some special cases that we can optimize globally by making a slight modification to the problem. Specifically, we will assume that the residuals are of the form

$$r_i(v) = \left\|\left(\frac{a_i^T v + \tilde{a}_i}{c_i^T v + \tilde{c}_i}, \frac{b_i^T v + \tilde{b}_i}{c_i^T v + \tilde{c}_i}\right)\right\|, \quad (2)$$

where $c_i^T v + \tilde{c}_i > 0$, that is, each coordinate is a quotient of affine functions. Instead of solving the least squares formulation we will consider the optimization problem

$$\min_v \max_i \quad r_i(v) \quad (3)$$
$$\text{s.t.} \quad c_i^T v + \tilde{c}_i > 0, \quad (4)$$

which can be solved globally optimally because of some convexity properties.

Below we give two examples of problems where the residuals are of the type (2).

Triangulation In this problem we know the camera matrices $P_i = [A_i \quad t_i]$ and image points x_i and want to find the scene point X. The residuals are of the form

$$r_i(X) = \left\|\left(x_i^1 - \frac{A_i^1 X + t_i^1}{A_i^3 X + t_i^3}, x_i^2 - \frac{A_i^2 X + t_i^2}{A_i^3 X + t_i^3}\right)\right\|. \quad (5)$$

To see that this expression is of the correct type (2) we can re write it as

$$\left\|\left(\frac{(x_i^1 A_i^3 - A_i^1)X + x_i^1 t_i^3 - t_i^1}{A_i^3 X + t_i^3}, \frac{(x_i^2 A_i^3 - A_i^2)X + x_i^2 t_i^3 - t_i^2}{A_i^3 X + t_i^3}\right)\right\|. \quad (6)$$

The constraint $A_i^3 X + t_i^3 > 0$ means that the scene point should be in front of the camera.

Resection In the resection problem we want to estimate the camera parameters A and t from scene points X_i and their projections (x_i^1, x_i^2). The residuals of this problem are

$$r_i(A,t) = \left\|\left(x_i^1 - \frac{A^1 X_i + t^1}{A^3 X_i + t^3}, x_i^2 - \frac{A^2 X_i + t^2}{A^3 X_i + t^3}\right)\right\|. \quad (7)$$

Structure from Motion with Known Camera Orientations If the camera orientations of all the cameras are known (in practice computed with some other method such as the one described in Section **??**) then we can solve for both the 3D points and the camera positions simultaneously. In this case the unknowns are X and t which gives residuals of the form

$$r_i(X,t) = \left\|\left(x_i^1 - \frac{A^1 X_i + t^1}{A^3 X_i + t^3}, x_i^2 - \frac{A^2 X_i + t^2}{A^3 X_i + t^3}\right)\right\|. \quad (8)$$

2 Convex Optimization

In this section we review some properties of convex functions and sets that will be useful for solving (3)-(4).

2.1 Convex Sets

A set $C \in \mathbb{R}^n$ is called convex if the line segment joining any two points in C is contained in C. That is, if $x, y \in C$ then $\lambda x + (1-\lambda)y \in C$ for all λ with $0 \leq \lambda \leq 1$.

We call a point x of the form

$$x = \sum_{i=1}^{n} \lambda_i x_i, \tag{9}$$

where $\sum_{i=1}^{n} \lambda_i = 1$, $0 \leq \lambda_i \leq 1$ and $x_i \in C$, a convex combination of the points $x_1, ..., x_n$. A convex set always contains every convex combination of its points. Furthermore, it can be shown that a set is convex only if it contains all its convex combinations. Figure 1 shows simple examples of the notions introduced.

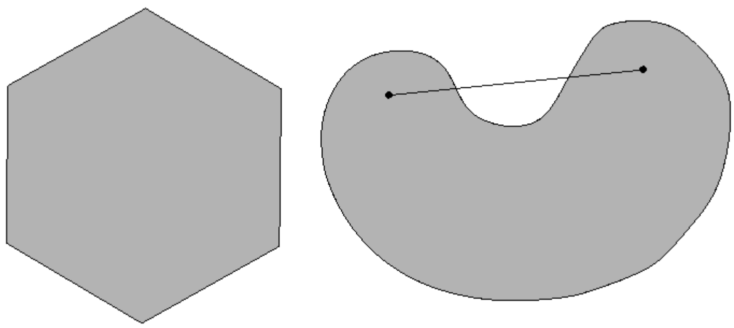

Figure 1: Left: A convex set. Right: A non-convex set.

Next we will state two special cases of convex sets that will be useful to us.

The halfspace. A halfspace is a set of the form

$$\{x \in \mathbb{R}^n; a^T x \leq b\}, \tag{10}$$

where $a \neq 0$, i.e., it is the solution set of a nontrivial affine inequality. The boundary of the half space is the hyperplane $\{x \in \mathbb{R}^n; a^T x = b\}$. It is straight forward to verify that these sets are convex.

The second order cone. The second order cone in $\mathbb{R}^{n+1}$ is the set

$$\{(x,t) \in \mathbb{R}^{n+1}; ||x|| \leq t\}. \tag{11}$$

From the general properties of norms it follows that the second order cone is a convex set in $\mathbb{R}^{n+1}$.

If $f : \mathbb{R}^m \mapsto \mathbb{R}^n$ is an affine mapping then the set $C' = \{x; f(x) \in C\}$ is convex in $\mathbb{R}^m$ if C is convex in $\mathbb{R}^n$. That is, convexity is preserved under affine mappings. When applied to the second order cone we get sets of the type

$$\{x; ||Ax + b||_2 \leq c^T x + d\}. \tag{12}$$

Convexity is also preserved under intersection. Thus a set C that is given by several of the constraints above (half spaces and cone-constraints) is a convex set.

2.2 Convex Functions

A function $f : C \mapsto \mathbb{R}$ is called convex if C is a convex set, and for all $x, y \in C$ and $0 \leq \lambda \leq 1$, we have

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y). \tag{13}$$

The geometric interpretation of this definition is that the line segment between the points $(x, f(x))$ and $(y, f(y))$ should lie above the graph of f. Figure 2 shows the geometric interpretation of the definition.

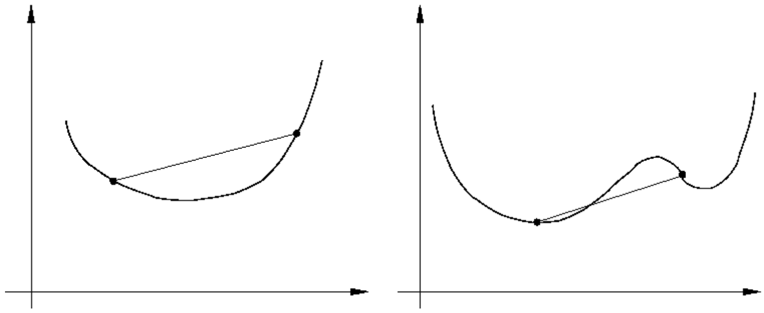

Figure 2: Left: Graph of a convex function. The line segment joining two points $(x, f(x))$ and $(y, f(y))$ lies above the graph. Right: Graph of a non-convex function.

Theorem 1. *If f is convex on a convex set C then any local minimum is also global.*

Proof. Asume that x is a local minimizer. Then there is a local neighborhood around x such that

$$f(x) \leq f(y) \tag{14}$$

for all y such that $\|x - y\| \leq \delta$. Suppose that x is not the global minimizer. Then there is an x^* such that

$$f(x^*) < f(x). \tag{15}$$

Since C is convex we can form the line segment between x^* and x and look at the values of f.

$$f(\lambda x^* + (1-\lambda)x) \leq \lambda f(x) + (1-\lambda)f(x^*) < \lambda f(x) + (1-\lambda)f(x) = f(x), \tag{16}$$

if $0 < \lambda \leq 1$. Now if we choose λ small enough such that $y = \lambda x^* + (1-\lambda)x$ fulfills $\|x - y\| \leq \delta$ we see that x cannot be a local minimizer since

$$f(x) > f(y). \tag{17}$$

□

3 Solving the Min-Max Problem

We now return to the min-max problem (3)-(4). By adding an extra variable γ we can rewrite the problem as

$$\min_{\gamma,v} \quad \gamma \tag{18}$$
$$\text{s.t.} \quad r_i(v) \leq \gamma, \quad \forall i. \tag{19}$$

Since we minimize γ and $\gamma \geq r_i(v)$ for all i, γ has to take the same value as the largest residual $\max_i r_i(v)$. Therefore the two formulations are equivalent. Since $c_i^T v + \tilde{c}_i > 0$ and we can write the problem as

$$\min_{\gamma,v} \quad \gamma \tag{20}$$
$$\text{s.t.} \quad \left\| \left(a_i^T v + \tilde{a}_i, b_i^T v + \tilde{b}_i \right) \right\| \leq \gamma (c_i^T v + \tilde{c}_i), \quad \forall i. \tag{21}$$

For a fixed γ the constraint (21) of the type (12) and therefore convex. Furthermore the intersection of all these constraints is also convex. This makes it possible to determine if there is a set of variables v that fulfill all the constraints for a given γ. Specifically, we can solve the convex program

$$\min_{s,v} \quad s \tag{22}$$
$$\text{s.t.} \quad \left\| \left(a_i^T v + \tilde{a}_i, b_i^T v + \tilde{b}_i \right) \right\| \leq \gamma(c_i^T v + \tilde{c}_i) + s, \quad \forall i. \tag{23}$$

If the optimal $s > 0$ then it is not possible to find a v that fulfills all the constraints at the same time. In this case we know that the current γ is smaller than $\min_v \max_i r_i(v)$. Otherwise if the optimal $s \leq 0$ we know that the current γ is larger than $\min_v \max_i r_i(v)$. This makes it possible to search for the $\min_v \max_i r_i(v)$ by solving a sequence of convex problems.

Below we outline a simple procedure, called bisection, that finds the optimal solution.

1. Let γ_l and γ_u be lower and upper bounds on the optimal error.
2. Check if there is a solution such that
$$r_i(v) \leq \frac{\gamma_u + \gamma_l}{2}, \quad \forall i$$
(convex optimization problem).
3. If there is set $\gamma_u = \frac{\gamma_u + \gamma_l}{2}$, otherwise set $\gamma_l = \frac{\gamma_u + \gamma_l}{2}$.
4. If $\gamma_u - \gamma_l > tol$ (some predefined tolerance) goto 2.

The result is an interval $[\gamma_l, \gamma_u]$ that is guaranteed to contain the optimal value.

4 Rotation Averaging

The methods that we described in the previous section can be used to solve the structure from motion problem if we have estimates for the camera rotations by optimizing over camera positions and 3D points. For this to be useful we need a method that can accurately estimate camera rotations which we will describe in this section.

For each pair of cameras that have a sufficient number of matches we can solve the (calibrated) relative pose problem (see Lecture 6). This gives us two cameras $P_1 = \begin{bmatrix} I & 0 \end{bmatrix}$ and $P_2 = \begin{bmatrix} R_{12} & t_{12} \end{bmatrix}$. The rotation R_{12} essentially tells us how we should rotate camera 1 to get camera 2. Similarly, if we solve the relative pose between cameras 2 and 3 we get the solution $P_2 = \begin{bmatrix} I & 0 \end{bmatrix}$ and $P_3 = \begin{bmatrix} R_{23} & t_{23} \end{bmatrix}$. Note that the solution to the relative pose problem is only defined up to a similarity transformation. That is, if we have one solution we may rotate translate and scale the solution to obtain a new one. Now let's say that the camera orientations are R_1,R_2 and R_3 in some common coordinate system. Then, by applying the rigid transformation $\begin{bmatrix} R_1^T & 0 \\ 0 & 1 \end{bmatrix}$ to the first solution we get

$$\begin{bmatrix} I & 0 \end{bmatrix} = \begin{bmatrix} R_1 & 0 \end{bmatrix} \begin{bmatrix} R_1^T & 0 \\ 0 & 1 \end{bmatrix} \tag{24}$$

and

$$\begin{bmatrix} R_{12} & t_{12} \end{bmatrix} = \begin{bmatrix} R_{12}R_1 & t_{12} \end{bmatrix} \begin{bmatrix} R_1^T & 0 \\ 0 & 1 \end{bmatrix} \tag{25}$$

That is $R_2 = R_{12}R_1$. Similarly by applying $\begin{bmatrix} R_2^T & 0 \\ 0 & 1 \end{bmatrix}$ to the second camera pair we get $R_3 = R_{23}R_2 = R_{23}R_{12}R_1$, which gives us estimates for the three camera orientations in the same coordinate system.

Now suppose that we also solve for the relative orientation between cameras 3 and 1 giving us the additional equation $R_1 = R_{31}R_3$. The system is now over-determined and due to noise it will in general not be possible to find an exact solution to all these equations at once. Least squares rotation averaging attempts to find the rotations that minimizes the problem

$$\min_{R_1,R_2,\ldots,R_n} \quad \|R_{ij}R_i - R_j\|^2 \tag{26}$$
$$\text{such that} \quad R_i^T R_i = I \quad \forall i = 1, \ldots, n. \tag{27}$$

This is a quadratic objective function with quadratic equallity constraints which in general results in a non-convex problem. In the following section we will describe how such problems can be addressed using Lagrangian duality.

5 Duality

Suppose that we have an a quadratic objective function $q(x) = x^T \tilde{A} x + 2b^T x + c$, where the variable $x \in \mathbb{R}^n$. By extending x to $\mathbb{R}^{n+1}$ and letting the extra coordinate be one we can write the objective as $q(x) = x^T A x$ where

$$A = \begin{bmatrix} \tilde{A} & b \\ b^T & c \end{bmatrix}. \tag{28}$$

Similarly suppose that we have quadratic functions $q_i(x) = x^T A_i x - 1$ and we want to solve

$$\min_x \quad q(x) \tag{29}$$
$$\text{such that} \quad q_i(x) - 1 = 0. \tag{30}$$

We will call this problem the primal problem. To eliminate the constraint we form the Lagrangian $L(x, \lambda) = q(x) + \sum_i \lambda_i (q_i(x) - 1)$ and consider

$$\min_x L(x, \lambda) \tag{31}$$

for a given value of λ. Let x^* be the minimizer of of (31). If x^* fulfills the constraints $q_i(x^*) = 0$ then we have found the minimizer of the original problem since $L(x, \lambda) = q(x)$ for all x that fulfill $q_i(x) = 0$. If the minimizer does not fulfill the constraints then we found a solution to that was better than all other points fulfilling the constraints. In both cases we have that $L(x^*, \lambda) < \min_x q(x)$ such that $q_i(x) = 0$. That is for each value of λ we get a lower bound on the optimal value of the primal problem. Lagrangian duality consists in trying to find the largest lower bound by solving

$$\max_\lambda \min_x L(x, \lambda). \tag{32}$$

We will refer to this as the dual problem. When both the objective and the constraints are quadratic it is possible to solve the inner minimization in closed from since $L(x, \lambda) = x^T (A + \sum_{i=1}^n \lambda_i A_i) x - \sum_{i=1}^n \lambda_i$. The minimum of this function is

$$L(x^*, \lambda) = \begin{cases} -\sum_{i=1}^n & \text{if } A + \sum_{i=1}^n \lambda_i A_i \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \tag{33}$$

Since the dual problem maximizes with respect to λ we do not need to consider the second case. The dual problem is then

$$\max_\lambda \quad -\sum_{i=1}^n \lambda_i \tag{34}$$
$$\text{such that} \quad A + \sum_{i=1}^n \lambda_i A_i \succeq 0. \tag{35}$$

This problem has a linear objective function and a convex constraint and is therefore convex and can be reliably solved with standard solvers.

For rotation averaging the dual problem has the from

$$\max_\Lambda \quad -\text{trace}(\Lambda) \tag{36}$$
$$\text{such that} \quad \Lambda - M \succeq 0, \tag{37}$$

where

$$M = \begin{bmatrix} 0 & R_{12} & \dots & R_{1n} \\ R_{21} & 0 & \dots & R_{2n} \\ \vdots & & \ddots & \vdots \\ R_{n1} & R_{n2} & \dots & 0 \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} \Lambda_1 & 0 & 0 & \dots \\ 0 & \Lambda_2 & 0 & \dots \\ 0 & 0 & \Lambda_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{38}$$

While the dual problem is in general only a lower bound on the primal problem it can be shown that for the rotation averaging problem it is tight if the solution has small enough error residuals. Specifically, if the rotation angle of the rotation $R_j^T R_{ij} R_i$ is smaller than $42.9°$ then the lower bound is tight.