MAE 546 Notes

Max Chien

Fall 2025

Contents

1 Preliminaries **3**
- 1.1 Motivations . . . 3
- 1.2 Definitions and Conventions . . . 5
- 1.3 Unconstrained Optimization . . . 5
- 1.4 Equality Constrained Optimization . . . 7
- 1.5 Mixed Constraint Mathematical Programs . . . 11

Definitions **14**

Introduction

Chapter 1

Preliminaries

1.1 Motivations

We denote by $\mathcal{A}$ the standard problem

$$\inf_{u \in \mathcal{U}} \left\{ J(u; t_0, t_f, X_0) = K(t_f, X_f) + \int_{t_0}^{t_f} L(s, X_s, u_s) \, \mathrm{d}s \right\}$$

where J is the objective function which we want to minimize, u is our control state from the admissible control $\mathcal{U}$, K is the terminal cost, L is the running cost, and the systeem is driven by a vector field f with

$$\mathrm{d}X_t = f(t, X_t, u_t) \, \mathrm{d}t$$

We may also need to satisfy equality constraints (like boundary conditions) and inequality constraints (like path constraints or bounds). If we impose regularity demands on any of the cost functions, solutions, or constraints, which will in turn change the conditions for solutions. We will also focus on finding local minima, though conditions like convexity can elevate these to global minima.

Example 1.1: Double Integrator Problem

Consider the minimum time problem where the cost function is given by

$$J(u; t_0, X_0) = \int_{t_0}^{t_f(u)} \mathrm{d}s = t_f(u) - t_0$$

where the dynamics are

$$\ddot{X}(t) = u(t)$$

and the system ends at time $t_f(u)$ when it is stopped, in other words

$$X_f = 0$$
$$\dot{X}_f = 0$$

In essence the goal is merely to stop at the origin as quickly as possible, within the admissible control set. Here we'll use $\mathcal{U} = \mathrm{PC}([t_0, \infty] \to [-1, 1])$, where PC denotes the set of piecewise continuous functions.

The solutions satisfy the "**bang bang principle**", where the optimal solution u^* takes values only on the vertices of the range; that is, its range is in $\{\pm 1\}$. It will be governed by a switching function φ and a **costate** or **adjoint** p^*, under

$$u^* = \begin{cases} 1, & \varphi(t; p^*) > 0 \\ -1, & \varphi(t; p^*) < 0 \\ \pm 1, & \varphi(t, p^*) = 0 \end{cases}$$

Solutions to this problem are known as **closed loop solutions**, meaning that the solution can be built over time by measuring the feedback output, as opposed to solving for the entire solution at once.

Example 1.2: Linear Quadratic Regulator

Consider the case of a lunar landar attempting to following a trajectory γ, but which has some error in its position (i.e. off course). We can compute a retargeting flight path $\delta\gamma$ using the linearization

$$\dot{\delta\gamma} \approx \nabla_\gamma f \delta\gamma + \nabla_u g \delta u$$

Here the cost function is given quadratically as

$$J(u; t_0, t_f, X_0) = \frac{1}{2} \int_{t_0}^{t_f} \langle X_s, Q(s) X_s \rangle + \langle u_s, R(s) u_s \rangle \, \mathrm{d}s$$

where Q, R are symmetric, Q is positive semidefinite, and R is positive definite. and the dynamics are

$$\mathrm{d}X_t = A(t) X_t \, \mathrm{d}t + B(t) u_t \, \mathrm{d}t$$

$A \in \mathbb{R}^{m \times m}, B \in \mathbb{R}^{m \times n}$, and the admissible control set is $C^1([t_0, t_f] \to \mathbb{R}^n)$. This is solved by

$$u_t^* = -R^{-1}(t) B^T(t) P(t) X_t^*$$

with P satisfying the **Riccati differential equation**

$$\dot{P}(t) = -P(t)A(t) - A^T(t)P(t) - Q(t) + P(t)B(t)R^{-1}(t)B^T(t)P(t)$$

and $P(t_f) = 0$.

In practice, control problems may be difficult or impossible to solve directly, so we may require transcription of the problem into a form amenable to numerical methods. This may be done directly, or first by deriving the necessary conditions through the costates.

There are a few methods for transcribing problems into a discretized form. **Shooting methods** involve transcription of only the control state, but the state process is still solved using the ODE involving f. For instance, if the admissible control states are $C^1([t_0, t_f] \to \mathbb{R})$, we might discretize $\mathcal{U}$ into four dimensions by replacing it with functions that take constant values on each of the four subintervals in $[t_0, t_f]$.

On the other hand, **collocation methods** transcribe both the control and state process at the same time.

1.2 Definitions and Conventions

We will denote a metric space by (M, d), and a topology by T. We assume all metric spaces are given the induced topology. For $x \in M$ a metric space, we denote the open ε-ball about x by $B(x, \varepsilon)$, and the closed ball by $\overline{B}(x, \varepsilon)$. The closure of a set A is denoted $\overline{A}$, its interior A^o, and its boundary ∂A.

Definition 1.1

If (X, T) is a topological space, then $x^* \in X$ is a local minimum for $f : X \to \mathbb{R}$ if there exists a neighborhood $A \in T$ of x^* where x^* minimizes f on A.

Definition 1.2

$C^k(\Omega, \mathbb{R})$ denotes the space of k times continuously differentiable functions from $\Omega \to \mathbb{R}$. $C_b(\Omega, \mathbb{R})$ is the space of such functions where all derivatives and the function are bounded.

1.3 Unconstrained Optimization

In this section we develop necessary and sufficient conditions for minima and strict minima on open sets in $\mathbb{R}^n$.

Proposition 1.1

If $x^* \in \Omega^o \subseteq \mathbb{R}^n$ is a local minimum for $f \in C^1(\Omega \to \mathbb{R})$, then

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

Theorem 1.2: Taylor's Formula with Remainder, Lagrange Form

Let $f \in C^{k+1}(\mathbb{R}, \mathbb{R})$. Let $x, x^* \in \mathbb{R}, \delta x = x - x^*$. Then there exists a point y strictly between x, x^* such that

$$f(x) = f(x^*) + f'(x^*)\delta x + \frac{1}{2!}f''(x^*)\delta x^2 + \ldots + \frac{1}{k!}f^{(k)}(x^*)\delta x^k + \frac{1}{(k+1)!}f^{(k+1)}(y)\delta x^{k+1}$$

Proposition 1.3

If $x^* \in \Omega^o \subseteq \mathbb{R}^n$ is a local minimum for $f \in C^2(\overline{\Omega}, \mathbb{R})$, then

$$\frac{\partial^2 f}{\partial x^2}|_{x^*} \geq 0$$

Definition 1.3

The **Hessian** of a function $f \in C^2(\Omega, \mathbb{R})$ at a point $x^* \in \Omega^o$ is

$$(\nabla_x^{\otimes^2} f|_{x^*})_{ij} = \partial_i \partial_j f$$

In particular the Hessian is symmetric.

Proposition 1.4

For $f \in C^2(\overline{\Omega}, \mathbb{R})$ and $\Omega \subseteq \mathbb{R}^n$, a sufficient condition for $x^* \in \Omega^o$ to be a strict local minimum of f is for

$$\nabla_x f|_{x^*} = 0$$
$$\nabla_x^{\otimes^2} f|_{x^*} > 0$$

(where the second line says the Hessian is positive definite.)

Proof. Since all the eigenvalues are positive, and the Hessian is symmetric, we write

$$\nabla_x^{\otimes^2} f|_{x^*} = Q\Lambda Q^T$$

such that

$$\left\langle \hat{q}_i, Q\Lambda Q^T \hat{q}_j \right\rangle = \delta_{ij}\lambda_i > 0$$

Then take $B(x^*, \varepsilon) \subseteq \Omega^o$ and define $g(\alpha, q) : [0, \varepsilon] \times S^{n-1} \to \mathbb{R}$ by $\alpha \times q \mapsto f(x^* + \alpha q)$. This gives the trace of f in the direction of q. Pick $q = \hat{q_1}$ and pick $\alpha \in (0, \varepsilon)$. By Taylor's theorem with remainder in α for $0 < \beta < \alpha$

$$g(\alpha, \hat{q}_1) = f(x^* + \alpha \hat{q}_1) = g|_0 + \partial_\alpha g|_0 \alpha + \frac{1}{2}\partial_\alpha^2 g(\hat{q}_1)|_\beta \alpha^2$$

By assumption, $\partial_\alpha g|_0 = \nabla_x f|_x^* \cdot \hat{q}_1 = 0$. So we see that

$$g(\alpha, \hat{q}_1) - g(0, \hat{q}_1) = \frac{1}{2}\partial_\alpha^2 g(\hat{q}_1)|_\beta \alpha^2$$

Assume $\alpha \ll 1$, so that

$$\text{sign}\left(\partial_\alpha^2 g(\hat{q}_1)|_\beta\right) = \text{sign}\left(\partial_\alpha^2 g(\hat{q}_1)|_0\right)$$

(possible since f is C^2). This shows that $f(x^* + \alpha\hat{q}_1) > f(x^*)$ for $0 < \alpha < \alpha_1^+$. We can repeat this work to show the same for $-\alpha_1^- < \alpha < 0$. We can also repeat this for the other eigenvalues. Finally set $\alpha^* = \min\{\alpha_i^+, \alpha_i^-\}$. It follows that

$$f(x^*) < f(y)$$

for any $y \in B(x^*, \alpha)$. □

Theorem 1.5: Taylor's Formula with Remainder, Peano Form

Let $f \in C^K(\mathbb{R}, \mathbb{R})$ and $x, x^* \in \mathbb{R}$, with $\delta x := x - x^*$. Then there exists $R_K : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \sum_{i=0}^{K} \frac{1}{i!} \partial_x^i f|_x^* \delta x^i + R_K(x)\delta x^K$$

such that $\lim_{x \to x^*} R_K(x) = 0$. For convenience we use asymptotic notation

$$f(x) = \sum_{i=0}^{K} \frac{1}{i!} \partial_x^i f|_x^* \delta x^i + o(\delta x^K)$$

Alternate Proof of 1.4. Use the Peano form to write

$$f(x^* + \alpha q) = f(x^*) + \frac{1}{2}\left\langle q, \nabla^{\otimes^2} f|_{x^*}, q\right\rangle \alpha^2 + o(\alpha^2, q)$$

For $q \in S^{n-1}$, define

$$h(q) = \sup\left\{\varepsilon > 0 : \alpha \in B(0,\varepsilon) \setminus \{0\} \implies \left|o(\alpha^2, q)\right| < \frac{1}{2}\left\langle q, \nabla_x^{\otimes^2} f|_{x^*} q\right\rangle \alpha^2\right\}$$

By compactness, h attains a minimum on S^{n-1}, so there exists ε^* such that the inequality is true on $B(0, \varepsilon^*) \setminus \{0\}$. □

1.4 Equality Constrained Optimization

Now we introduce equality constraints to study more interesting sets over which we may optimize. For $m \leq n$, define a set of constraints

$$\left\{h_i \in C^1(\mathbb{R}^n, \mathbb{R})\right\}_{i=1}^m$$

and define the collective zero locus

$$M = \bigcap_i \{h_i = 0\}$$

We will always assume that our constraints are nondegenerate, so that $M \neq \varnothing$.

Definition 1.4

A **regular point** is an element $q \in M$ such that the gradients

$$\{\nabla_x h_i|_q\}_i$$

are linearly independent. Note that if any gradient is zero, then q is not regular.

Definition 1.5

Let $h \in C^1(\Omega, \mathbb{R}^m)$, $\Omega \subseteq \mathbb{R}^n$. Then the **Jacobian** of h at $q \in \Omega^o$ is

$$(\nabla_x h|_q)_{ij} = \frac{\partial h_i}{\partial x_j}|_q = \begin{bmatrix} \nabla h_1^T \\ \vdots \\ \nabla h_m^T \end{bmatrix}$$

If the Jacobian is full rank, that is $\mathrm{rank}(\nabla_x h|_q) = \min(m, n)$, then q is a regular point. We define the tangent space in two equivalent ways:

Definition 1.6: Tangent Space, Geometric

Let $q \in M = M_k$ be a point on a k-dimensional surface. The **tangent space** to M at q, denoted T_qM, is the vector space isomorphic to $\mathbb{R}^k$ defined iby

$$T_qM \coloneqq \left\{(q, y) \in M \times \mathbb{R}^k : \langle \nabla_x h_i, y\rangle = 0 \quad \forall i\right\}$$

Definition 1.7: Tangent Space, Curves

Consider the family of curves $\left\{\psi_\lambda \in C^1((-1,1), M_K)\right\}_{\lambda\in\Lambda}$ such that $\psi_\alpha(0) = q$. Let $f \in C^1(M_K, \mathbb{R})$. Then by the chain rule,

$$\partial_\alpha(f \circ \psi_\lambda)|_0 = \langle \nabla_x f|_q, \partial_\alpha \psi_\lambda|_0\rangle$$

In particular for $f = h_i$, $h_i(\psi_\lambda(\alpha)) \equiv 0$, so

$$\langle \nabla_x h_i|_q, \partial_\alpha \psi_\lambda(0)\rangle = 0$$

This is the same inner product condition as the geometric definition, so we can just define the tangent space to be the collection of $\partial_\alpha \psi_\lambda(0)$, endowed with vector space structure and equivalence via curve equivalence.

Now we give necessary conditions on optimization on equality hypersurfaces.

Proposition 1.6

Let $M = M_k \subseteq \mathbb{R}^n$ and $k = n - m$, with M defined by $(h_i)_{i=1}^m$. If $x^* \in M$ is a minimum of $f \in C^1(M, \mathbb{R})$ and x^* is a regular point, then there exists $\lambda \in \mathbb{R}^m$ such that

$$0 = \nabla_x f|_{x^*} + \nabla_x h|_{x^*} \lambda$$

In other words, f is linearly dependent with the gradients of the constraints.

Proof. Since x^* is a regular point, we can form a basis of $\mathbb{R}^n$ given by the m gradients $(\nabla_x h_i|_{x^*})_{i=1}^m$ and a basis of $T_{x^*}M$ (say, $(\partial_\alpha \psi_j(0))_{j=1}^k$ for some ψ_j). Thus we can write $\nabla_x f|_{x^*}$ as

$$\nabla_x f|_{x^*} = \sum_{i=1}^m \langle \nabla_x f|_{x^*}, \nabla_x h_i|_{x^*} \rangle \nabla_x h_i|_{x^*} + \sum_{j=1}^k \langle \nabla_x f|_{x^*}, \partial_\alpha \psi_j(0) \rangle \partial_\alpha \psi_j(0)$$

For any ψ_j, write $g_j = f \circ \psi_j$. Then $g \equiv 0$ since $f = 0$ on M_k, so

$$0 = \partial_\alpha|_0 = \langle \nabla_x f|_{x^*}, \partial_\alpha \psi|_0 \rangle$$

So $\nabla_x f|_{x^*}$ is a linear combination of the $\nabla_x h_i|_{x^*}$. $\square$

The above proof is essentially a statement that the method of **Lagrange multipliers** works.

Analytic Proof. This proof works for $m = 1$. Let $d_1, d_2 \in \mathbb{R}^n$ and define $F : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$F(\alpha_1, \alpha_2) = (f(x^* + \alpha_1 d_1 + \alpha_2 d_2), h(x^* + \alpha_1 d_1 + \alpha_2 d_2))$$

In particular $F(0, 0) = (f(x^*), 0)$. Now consider the matrix

$$\nabla F|_{(0,0)} = \begin{bmatrix} \langle \nabla f, d_1 \rangle & \langle \nabla f, d_2 \rangle \\ \langle \nabla h, d_1 \rangle & \langle \nabla h, d_2 \rangle \end{bmatrix}$$

Suppose the rank of this matrix is 2. Then F is locally invertible at x^*. So there is an open neighborhood around $(f(x^*), 0)$ where F is invertible, and by passing through the inverse map, there is (α_1, α_2) such that

$$\pi_1 \circ F(\alpha_1, \alpha_2) = f(x^* + \alpha_1 d_1 + \alpha_2 d_2) < f(x^*)$$
$$\pi_2 \circ F(\alpha_1, \alpha_2) = 0$$

But this is a contradiction. So ∇F is not full rank. Since x^* is a regular point, we can choose d_1 such that $\langle \nabla h, d_1 \rangle \neq 0$. Let d_2 be arbitary, and define

$$\lambda^* = -\frac{\langle \nabla f, d_1 \rangle}{\langle \nabla h, d_1 \rangle}$$

Now, ∇F has rank exactly 1, so the columns are proportional. This means there is β such that

$$\langle \nabla h, d_1 \rangle = \frac{1}{\beta}\langle \nabla h, d_2 \rangle$$
$$\langle \nabla f, d_1 \rangle = \frac{1}{\beta}\langle \nabla f, d_2 \rangle$$

Then

$$\langle \nabla f, d_2 \rangle = \beta(\langle \nabla f, d_1 \rangle) = \beta(-\lambda\langle h, d_1 \rangle) = -\lambda\langle \nabla h, d_2 \rangle \implies \langle \nabla f + \lambda \nabla h, d_2 \rangle = 0$$

Since d_2 is arbitrary,

$$\nabla f + \lambda \nabla h = 0$$

$\square$

Definition 1.8

Let $h : \mathbb{R}^n \to \mathbb{R}^m$ and define the **augmented Lagrangian cost function** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$\mathcal{L}(x, \lambda) = f(x) + \langle \lambda, h(x) \rangle$$

Corollary 1.7

In the same setup as the previous theorem, there is $\lambda^* \in \mathbb{R}^m$ such that

$$\nabla_x \mathcal{L}|_{(x^*,\lambda^*)} = \nabla_x f|_{x^*} + \nabla h_{x^*}^T \lambda^* = 0$$

and

$$\nabla_\lambda \mathcal{L}|_{(x^*,\lambda^*)} = h(x^*) = 0$$

Essentially, the Lagrangian extends our constrained optimization to a higher dimension space, on which we may perform unconstrained optimization (so long as the minimum is regular). Thus the necessary and sufficient conditions look very similar to the unconstrained case.

Theorem 1.8: Second Order Necessary Condition

Let M be the zero locus of $h : \mathbb{R}^n \to \mathbb{R}^m$, with $h_i \in C^2(\mathbb{R}^n, \mathbb{R})$, $f \in C^2(M, \mathbb{R})$, and let $\mathcal{L}$ be the augmented Lagrangian. Then the Hessian of the augmented Lagrangian with respect to x is

$$\nabla_x^{\otimes^2} \mathcal{L}|_{(x,\lambda)} = \nabla_x^{\otimes^2} f|_{(x,\lambda)} + \sum_i \lambda_i \nabla_x^{\otimes^2} h_i|_{(x,\lambda)}$$

Moreover, if x^* is a minimum of f and a regular point of M, then there exists $\lambda^* \in \mathbb{R}^m$ such that $\nabla_x^{\otimes^2} \mathcal{L}_{(x^*,\lambda^*)}$ is positive semidefinite.

Theorem 1.9: Second Order Sufficient Condition

If

$$\nabla_x \mathcal{L}|_{(x^*,\lambda^*)} = 0 \in \mathbb{R}^n$$
$$\nabla_\lambda \mathcal{L}|_{(x^*,\lambda^*)} = 0 \in \mathbb{R}^M$$

and $\nabla_x^{\otimes^2} \mathcal{L}|_{(x^*,\lambda^*)}$ is positive definite, and moreover x^* is regular, then x^* is a strict local minimum of f.

1.5 Mixed Constraint Mathematical Programs

Definition 1.9

A **mixed constraint mathematical program** is a problem of the form of finding

$$\inf_{\mathbb{R}^n} f$$

subject to the constraints

$$\begin{aligned} h_e(x) &= 0, \quad e \in E \\ c_i(x) &\leq 0, \quad i \in I \end{aligned}$$

with $|E| = m < n$ and $|I| \in \mathbb{N}$. When f, h, c are all linear functions, this is called a **linear program** (LP); when f is quadratic and h, c are linear, this is a **quadratic program** (QP). If f, h, c are all convex, then it is called a **convex program** (CVP). Most generally, this can be called a **nonlinear program** (NLP).

While solving NLPs, it is often helpful to break it into sequential programs of simpler type, like QPs or CVPs. For instance, **sequential quadratic programs** (SQP) involve a method similar to gradient descent, but by solving a QP at every step, since we know f locally looks like a QP at a minimum.

Definition 1.10

Let $A \subseteq \mathbb{R}^m$ be convex, and let $f : A \to \mathbb{R}$. Define the **epigraph** of f by

$$B = \{(x, y) : x \in A, y \geq f(x)\} \subseteq A \times \mathbb{R}$$

f is said to be **convex** if B is convex in $\mathbb{R}^{m+1}$. Equivalently, f is said to be convex if it is continuous and for $x, y \in A, t \in [0, 1]$,

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y)$$

Definition 1.11

A set $\mathcal{C} \subseteq \mathbb{R}^n$ is called a **cone** if for all $x \in \mathcal{C}$, $t > 0$, $tx \in \mathcal{C}$.

Definition 1.12

For a mixed constraint program with equality constraints $h_e, e \in E$ and inequality constraints $c_i, i \in I$, the **feasible set** is the set

$$\Omega = \{x \in \mathbb{R}^n : c_i(x) = 0, h_e(x) = 0\}$$

The **active set** at a point $x \in \Omega$ is the set of indices for which x achieves equality; that is,

$$A(x) = \{i \in I : c_i(x) = 0\} \sqcup E$$

Example 1.3

Suppose $f \in C^1(\mathbb{R}^n, \mathbb{R})$ and let $c \in C^1$ be the only inequality constraint. Let $x \in \Omega$ be a point in the feasible set. Let us try to find $q \in S^{n-1}, \alpha > 0$ such that $x + \alpha q \in \Omega$ and $f(x + \alpha q) < f(x)$.

If $c(x) < 0$ then $A(x) = \varnothing$, otherwise if $c(x) = 0$ then $A(x) = \{1\}$. In the first case this locally just looks like unconstrained optimization and we are done by our previous work, seting $q = -\nabla_x f|_x$.

Otherwise, we want to have $\langle \nabla f|_x, q\rangle < 0$ and $c(x + \alpha q) \leq 0$. Suppose such q, α exist. Applying the mean value theorem to c, there is $0 < \beta < \alpha$ such that

$$c(x + \alpha q) = c(x) + \alpha \langle \nabla c|_{x+\beta q}, q\rangle = \alpha \langle \nabla c|_{x+\beta q}, q\rangle \leq 0$$

Let α be small enough such that for all $\beta < \alpha$,

$$\operatorname{sign}(\langle \nabla c|_x, q\rangle) = \operatorname{sign}(\langle \nabla c|_{x+\beta q}, q\rangle)$$

So in particular we have

$$\langle \nabla f|_x, q\rangle < 0, \qquad \langle \nabla c|_x, q\rangle \leq 0$$

As a result, this cannot happen (which occurs at minima) if

$$\langle \nabla f|_x, q\rangle = -\lambda \nabla c|_x$$

for some $\lambda \geq 0$. A concise way to express conditions for this under both cases of $c(x)$ is that there exists $\lambda \geq 0$ such that

$$\nabla_x \mathcal{L}|_{(x,\lambda)} = 0$$
$$\lambda c(x) = 0$$

The second of these conditions is called the **complementarity condition**.

Definition 1.13

We say that the **linear independence constraint qualification** (LICQ) holds at a point $a \in \Omega$ if

$$\operatorname{span}\{\nabla c_i : i \in A(x)\} = \mathbb{R}^{|A(x)|}$$

Proposition 1.10

Suppose x^* is a minimum of f and the LICQ holds at x^*. Then there exists $\lambda^* \in \mathbb{R}^{|E \sqcup I|}$ such that

$$\nabla_x \mathcal{L}|_{(x^*,\lambda^*)} = 0$$

and $\lambda_i^* \geq 0$ for all $i \in I$ with

$$\lambda_i^* c_i(x) = 0, \quad i \in I$$

Definitions

active set, 12
adjoint, 4
augmented Lagrangian cost function, 10

bang bang principle, 4

closed loop solutions, 4
collocation methods, 5
complementarity condition, 12
cone, 12
convex, 11
convex program, 11
costate, 4

epigraph, 11

feasible set, 12

Hessian, 6

Jacobian, 8

Lagrange multipliers, 9
linear independence constraint qualification, 13
linear program, 11

mixed constraint mathematical program, 11

nonlinear program, 11

quadratic program, 11

regular point, 8
Riccati differential equation, 4

sequential quadratic programs, 11
Shooting methods, 4

tangent space, 8