Day 4: Parameter Estimation and Prediction in Epidemic Models

Danillo Barros de Souza

Basque Center for Applied Mathematics

January 22, 2026

Outline

SIR Model Overview

The population is divided into three compartments:

- $S(t)$: Susceptible individuals
- $I(t)$: Infectious individuals
- $R(t)$: Removed (recovered or deceased)

Total population:

$$S(t) + I(t) + R(t) = N \quad \text{(constant)}$$

SIR Model Equations

The deterministic SIR model is given by:

$$\begin{aligned}\frac{dS}{dt} &= -\beta\frac{SI}{N},\\ \frac{dI}{dt} &= \beta\frac{SI}{N} - \gamma I,\\ \frac{dR}{dt} &= \gamma I.\end{aligned}$$

- β: transmission rate
- γ: recovery rate

Why Estimate Parameters?

In real epidemics:

- β and γ are unknown
- We observe data: infected cases over time
- Parameters must be estimated from data

Goal:

Find β, γ such that model fits observed data

Observed Data

Assume we observe infected individuals at discrete times:

$$\{(t_k, I_k^{\mathrm{obs}})\}_{k=1}^{M}$$

The model prediction:

$$I_k^{\mathrm{mod}}(\beta, \gamma) = I(t_k; \beta, \gamma)$$

obtained by numerically solving the SIR system.

Least Squares Error

Define the least squares objective function:

$$J(\beta, \gamma) = \sum_{k=1}^{M} \left(I_k^{\text{mod}}(\beta, \gamma) - I_k^{\text{obs}} \right)^2$$

Parameter estimation problem:

$$(\hat{\beta}, \hat{\gamma}) = \arg\min_{\beta, \gamma} J(\beta, \gamma)$$

Interpretation of Least Squares

- Penalizes large deviations between model and data
- Assumes measurement noise is approximately Gaussian
- Widely used due to simplicity and interpretability

Optimization is typically done numerically:

- Gradient-based methods
- Grid search (for teaching purposes)

Using Estimated Parameters

Once parameters are estimated:

$$\beta = \hat{\beta}, \quad \gamma = \hat{\gamma}$$

We solve the SIR system forward in time to obtain:

$$I(t) \quad \text{for future } t$$

This allows epidemic forecasting.

New Infected Individuals

The number of **new infections per unit time** is:

$$\text{New infections}(t) = \beta \frac{S(t)I(t)}{N}$$

This quantity:

- Peaks before $I(t)$ peaks
- Is useful for public health planning

Definition of R_0

The basic reproduction number:

$$R_0 = \frac{\beta}{\gamma}$$

Interpretation:

- Average number of secondary infections caused by one infected individual
- Assumes a fully susceptible population

Epidemic Threshold

- If $R_0 > 1$: epidemic outbreak occurs
- If $R_0 < 1$: disease dies out

Using estimated parameters:

$$\hat{R}_0 = \frac{\hat{\beta}}{\hat{\gamma}}$$

Provides a quantitative measure of epidemic severity.

Summary

- SIR model describes disease dynamics with ODEs
- Least squares is used to estimate parameters from data
- Estimated model enables prediction of future infections
- R_0 summarizes epidemic potential

These tools form the foundation of modern epidemiological modeling.