Calculus Skills for Estimation I

PLSC 30700 — Pre-Midterm Review

Robert Gulotty
University of Chicago

Winter 2026

This sheet collects the calculus skills used in Lectures 1–7 (before the midterm). Linear algebra skills are covered in a separate document. Each section lists the rules you need and where they appear in the course. If any of these are unfamiliar, review them before the relevant lecture.

Contents

1 Derivative Rules (Single Variable)

These rules are the building blocks for everything that follows.

1.1 Power Rule

$$\frac{d}{dx}x^n = nx^{n-1}$$

Where it appears:

- **Day 1:** Minimizing MSE. $\frac{d}{d\mu}[\mu^2] = 2\mu$, used to show $\mathbb{E}[Y]$ minimizes $\mathbb{E}[(Y-\mu)^2]$.
- **Day 7 (MLE):** Maximizing the binomial likelihood $p^4(1-p)^6$.

1.2 Product Rule

$$\frac{d}{dx}[f(x) \cdot g(x)] = f'(x)g(x) + f(x)g'(x)$$

Where it appears:

- **Day 7:** Differentiating the binomial likelihood $p^4(1-p)^6$:

$$\frac{d}{dp}[p^4(1-p)^6] = 4p^3(1-p)^6 - 6p^4(1-p)^5$$

1.3 Chain Rule

$$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x)$$

Where it appears:

- **Day 7:** Differentiating $\log f(x|\theta)$ when f involves compositions (e.g., $\exp(-x^2/(2\theta))$).
- **Day 7:** MLE invariance: if $\hat{\theta}$ is the MLE of θ, then $h(\hat{\theta})$ is the MLE of $h(\theta)$.

1.4 Derivative of $\log x$

$$\frac{d}{dx}\log x = \frac{1}{x}, \qquad \frac{d}{dx}\log(g(x)) = \frac{g'(x)}{g(x)}$$

Where it appears:

- **Day 7:** Log-likelihood of the exponential: $\frac{d}{d\lambda}[-n\log\lambda] = -n/\lambda$.
- **Day 7:** Log-likelihood of the normal: $\frac{d}{d\sigma^2}[-\frac{n}{2}\log(\sigma^2)] = -\frac{n}{2\sigma^2}$.

1.5 Derivative of e^x and $e^{g(x)}$

$$\frac{d}{dx}e^x = e^x, \qquad \frac{d}{dx}e^{g(x)} = g'(x) \cdot e^{g(x)}$$

Where it appears:

- **Day 1:** The normal density $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.
- **Day 7:** Exponential family likelihoods, normal log-likelihood.

1.6 Derivative of $1/x$

$$\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}, \qquad \text{equivalently, } \frac{d}{dx}x^{-1} = -x^{-2}$$

Where it appears:

- **Day 7:** MLE for exponential distribution: $\frac{d}{d\lambda}\left[-\frac{\sum X_i}{\lambda}\right] = \frac{\sum X_i}{\lambda^2}$.

1.7 Second Derivatives and Second-Order Conditions

If $f'(x^*) = 0$ and $f''(x^*) > 0$, then x^* is a local minimum.
If $f'(x^*) = 0$ and $f''(x^*) < 0$, then x^* is a local maximum.

Where it appears:

- **Day 1:** Verifying that $\mu^* = \mathbb{E}[Y]$ minimizes MSE.
- **Day 3:** The Hessian of the SSE is $2\boldsymbol{X'X}$, which is positive definite $\Rightarrow$ unique minimum.
- **Day 7:** Verifying that the MLE is a maximum of the likelihood.

2 Integration

2.1 Definite Integrals and the Power Rule for Integration

$$\int x^n \, dx = \frac{x^{n+1}}{n+1} + C \quad (n \neq -1)$$

$$\int_a^b f(x)\, dx = F(b) - F(a) \quad \text{where } F'(x) = f(x)$$

Where it appears:

- **Day 1:** Computing expectations: $\mathbb{E}[Y] = \int y\, f(y)\, dy$.
- **Day 1:** Marginal density from joint: $f_X(x) = \int_0^{1-x} 6(1-x-y)\, dy$.
  - Requires integrating $\int (1-x)y - y^2/2$ and evaluating at limits.
- **Day 1:** Computing the CEF: $\mathbb{E}[Y|X=x] = \int y \cdot f_{Y|X}(y|x)\, dy$.

2.2 Double Integrals

$$\iint f(x,y)\,dx\,dy = \int \left[\int f(x,y)\,dy\right] dx$$

Integrate the inner variable first, treating the outer variable as a constant.

Where it appears:

- **Day 1:** Verifying a joint density integrates to 1: $\int_0^1 \int_0^{1-x} 6(1-x-y)\,dy\,dx = 1$.
- **Day 1:** Law of iterated expectations (proof): $\mathbb{E}[Y] = \int\int y\,f(x,y)\,dy\,dx$.

3 Partial Derivatives

3.1 Definition and Notation

$$\frac{\partial f(x,y)}{\partial x} = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}$$

Differentiate with respect to one variable, holding all others constant.

Where it appears:

- **Day 1:** Joint density from CDF: $f(x,y) = \frac{\partial^2}{\partial x\,\partial y} F(x,y)$.
- **Day 2:** Solving for BLP: take $\frac{\partial}{\partial a}$ and $\frac{\partial}{\partial b}$ of $\mathbb{E}[(Y-(a+bX))^2]$.
- **Day 2:** Marginal effects with interactions: $\frac{\partial \mathbb{E}[y|X]}{\partial x_i} = \beta_1 + \gamma z_i$.
- **Day 7:** MLE for normal regression: separate FOCs for β and σ^2.

4 Differentiating Through Expectations and Sums

$$\frac{d}{d\theta}\mathbb{E}[g(X,\theta)] = \mathbb{E}\left[\frac{\partial}{\partial\theta} g(X,\theta)\right] \quad \text{(under regularity conditions)}$$

$$\frac{d}{d\theta}\sum_{i=1}^{n} g(x_i,\theta) = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} g(x_i,\theta)$$

Where it appears:

- **Day 1:** $\frac{d}{d\mu}\mathbb{E}[(Y-\mu)^2] = \mathbb{E}\left[\frac{d}{d\mu}(Y-\mu)^2\right] = -2\mathbb{E}[Y] + 2\mu$.
- **Day 2:** FOC for BLP: $\frac{d}{da}\mathbb{E}[(Y-a-bX)^2] = -2\mathbb{E}[Y-a-bX]$.
- **Day 3:** FOC for OLS: $\frac{d}{d\beta}\sum(Y_i - \boldsymbol{x}_i'\beta)^2 = -2\sum \boldsymbol{x}_i(Y_i - \boldsymbol{x}_i'\beta)$.
- **Day 7:** Score of the log-likelihood: $\frac{d}{d\theta}\sum \log f(X_i|\theta) = \sum \frac{d}{d\theta}\log f(X_i|\theta)$.

5 Gradients (Vector Calculus)

5.1 Gradient of a Scalar Function

If $f(\boldsymbol{a})$ is a scalar function of a vector $\boldsymbol{a} \in \mathbb{R}^k$:

$$\nabla_{\boldsymbol{a}} f(\boldsymbol{a}) = \begin{pmatrix} \partial f/\partial a_1 \\ \vdots \\ \partial f/\partial a_k \end{pmatrix}$$

Setting $\nabla_{\boldsymbol{a}} f = \boldsymbol{0}$ gives the first-order conditions.

5.2 Key Formulas for Matrix Calculus

$$\begin{aligned} \nabla_{\boldsymbol{a}}(\boldsymbol{z}'\boldsymbol{a}) &= \boldsymbol{z} && \text{(derivative of a linear form)} \\ \nabla_{\boldsymbol{a}}(\boldsymbol{a}'\boldsymbol{Z}\boldsymbol{a}) &= (\boldsymbol{Z}+\boldsymbol{Z}')\boldsymbol{a} = 2\boldsymbol{Z}\boldsymbol{a} \text{ if } \boldsymbol{Z} \text{ symmetric} && \text{(derivative of a quadratic form)} \end{aligned}$$

Where it appears:

- **Day 1:** Deriving OLS from vector calculus.
- **Day 2:** Multivariate BLP: $\frac{\partial}{\partial \beta} S(\beta) = -2\mathbb{E}[XY] + 2\mathbb{E}[XX']\beta$.
- **Day 3:** Full OLS derivation:

$$\begin{aligned} \nabla_\beta[\boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\beta + \beta'\boldsymbol{X}'\boldsymbol{X}\beta] &= -2\boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X}\beta = \boldsymbol{0} \\ \Rightarrow \quad \hat{\beta} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \end{aligned}$$

- **Day 3 (SOC):** The Hessian is $2\boldsymbol{X}'\boldsymbol{X}$, which is positive definite under the rank condition.

6 Optimization (Putting It All Together)

Recipe for unconstrained optimization:

1. Write the objective function $f(\theta)$.
2. Compute the first derivative(s) and set equal to zero: $f'(\theta^*) = 0$ (FOC).
3. Solve for the critical point(s) θ^*.
4. Check the second-order condition: $f''(\theta^*) > 0$ means minimum, $f''(\theta^*) < 0$ means maximum.

This recipe is used in every major derivation:

Lecture	**Objective**	**Result**	**Min/Max**
Day 1	$\mathbb{E}[(Y-\mu)^2]$	$\mu^* = \mathbb{E}[Y]$	Min
Day 2	$\mathbb{E}[(Y-a-bX)^2]$	$b = \text{Cov}(X,Y)/\text{Var}(X)$	Min
Day 3	$(\boldsymbol{y}-\boldsymbol{X}\beta)'(\boldsymbol{y}-\boldsymbol{X}\beta)$	$\hat{\beta} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'y}$	Min
Day 7	$\mathcal{L}_n(\theta) = \prod f(X_i\|\theta)$	$\hat{\theta}_{MLE}$	Max
Day 7	Normal log-lik for (β, σ^2)	$\hat{\beta}_{MLE} = \hat{\beta}_{OLS}$	Max

7 Summary: Skills by Lecture Day

Lecture	**Calculus Skills Used**
Day 1 (Review)	Definite integrals (expectations, marginals, CEF); double integrals; partial derivatives; power rule; gradient and quadratic form derivatives; FOC/SOC
Day 2 (CEF/BLP)	Differentiating through expectations; partial derivatives for multi-parameter optimization; marginal effects as partial derivatives of regression function
Day 3 (OLS)	Matrix calculus (linear and quadratic forms); FOC in matrix notation; SOC via Hessian; differentiating sums
Day 4 (Sensitivity)	No new calculus (algebraic manipulation of OLS formulas)
Day 5 (GLS)	No new calculus (variance formulas, matrix algebra)
Day 6 (Heteroskedasticity)	No new calculus (variance estimator properties)
Day 7 (MLE)	Product rule; chain rule; $\frac{d}{dx}\log x$; $\frac{d}{dx}e^{g(x)}$; $\frac{d}{dx}x^{-1}$; second derivatives of log-likelihoods; score and Hessian; optimization of likelihoods