UJEP

Department of Mathematics

Mathematics for Data Engineers

KMA/MDI

Part II

Differential calculus
and optimization

Lectures 5–7

Yaroslav Bazaikin

17 October 2026

Preface

This part of *Mathematics for Data Engineers*, KMA/MDI, contains Lectures 5–7: differentiation in several variables, unconstrained optimization, and equality constraints. It continues Part I, using its Euclidean geometry, Gram matrices, and spectral description of quadratic forms.

We assume one-variable differentiation of polynomials, exponentials, and trigonometric functions. Each lecture retains six review questions and eight exercises, including two extensions marked with an asterisk. Sections labelled *Further perspective* can be read after the core.

The route through this part

Lecture	Mathematical tool and application
5	Linear derivatives and gradients: sensitivity, uncertainty, and informative measurements.
6	Taylor models and Hessians: classifying stationary points and fitting parameters.
7	Tangent spaces and multipliers: exact requirements, allocation, and the cost of a constraint.

The final bridge replaces a finite sum of squared residuals by an integral. The same Gram matrix argument then leads to approximation by functions and, later, Fourier coefficients. The Jacobian introduced in Lecture 5 also supplies the derivative matrix used for changes of variables in Part III.

Starting point and notation

We retain the temperature-channel calibration from revised Part I. At temperatures $T = 20, 25, 30\,^{\circ}\mathrm{C}$, the normalized input $t = (T - 25)/5$ has values $-1, 0, 1$, and the measured outputs are $1, 2, 4$ volts. A linear model predicts $a + bt$. In numerical voltage units,

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} a \\ b \end{pmatrix}, \qquad \boldsymbol{y} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

The *offset* a is the predicted output at $25\,^{\circ}\mathrm{C}$, where $t = 0$. The *slope* b is the output change per unit of t, hence per $5\,^{\circ}\mathrm{C}$. The data $A, \boldsymbol{y}$ stay fixed while we fit the parameters $\boldsymbol{c}$.

Throughout both revised parts, a residual means *observed minus predicted*:

$$\boldsymbol{r}(\boldsymbol{c}) = \boldsymbol{y} - A\boldsymbol{c}, \qquad J(a, b) = \|\boldsymbol{r}(\boldsymbol{c})\|^2 = (1 - a + b)^2 + (2 - a)^2 + (4 - a - b)^2. \tag{1}$$

There is no factor $1/2$ in J. Part I obtained

$$\boldsymbol{c}_* = (7/3, 3/2)^{\mathsf{T}}, \qquad A\boldsymbol{c}_* = (5/6, 7/3, 23/6)^{\mathsf{T}}, \qquad J(\boldsymbol{c}_*) = 1/6.$$

We will recover this optimum by differentiation and examine what changes when the endpoint prediction $a + b$ is prescribed.

Vectors are columns; $^{\mathsf{T}}$ means transpose and $\|\cdot\|$ is the Euclidean norm. A displacement is $\boldsymbol{h}$, a unit direction is $\boldsymbol{d}$, and a path is $\boldsymbol{x}(s)$. For a scalar function, Df is a linear map represented by a row, while its gradient ∇f is a column. The Hessian is H_f. For a vector-valued function F, DF is its Jacobian matrix.

Contents

Preface **i**

Starting point and notation **ii**

5 Gradients and local sensitivity **1**
5.1 Inputs, domains, and partial derivatives . . . 1
5.2 The gradient and the total derivative . . . 1
5.3 Directions and the steepest local change . . . 2
5.4 A moving input and the chain rule . . . 3
5.5 Finite changes and uncertainty . . . 4
5.6 Further perspective: amplitude and phase sensitivities . . . 5
5.7 Questions and exercises . . . 6

6 Curvature and unconstrained optima **7**
6.1 The Hessian and the second-order model . . . 7
6.2 Stationary points and necessary conditions . . . 8
6.3 What the Hessian can decide . . . 9
6.4 Least squares: curvature from the design matrix . . . 10
6.5 Convexity: when a local optimum is global . . . 11
6.6 Further perspective: choosing an optimization step . . . 11
6.7 Questions and exercises . . . 13

7 Equality-constrained optimization **14**
7.1 Feasible points and tangent directions . . . 14
7.2 Why multipliers appear . . . 14
7.3 A constraint changes the fitted optimum . . . 15
7.4 Allocation and equal marginal costs . . . 16
7.5 Curved constraints and global comparison . . . 17
7.6 Boundaries and singular constraints . . . 18
7.7 Further perspective: the meaning of a multiplier . . . 19
7.8 Questions and exercises . . . 21

From parameter geometry to function geometry **22**

Lecture 5

Gradients and local sensitivity

A measurement model maps several inputs to an output. Its derivative separates the sensitivities of the model from a particular choice of input changes.

5.1 Inputs, domains, and partial derivatives

Let $U \subseteq \mathbb{R}^n$ be open and $f : U \to \mathbb{R}$. Openness means that each $\boldsymbol{p} \in U$ admits all sufficiently small displacements. Restrictions such as positive amplitudes or nonzero denominators determine where a formula is defined. Additional feasibility constraints will be treated separately in Lecture 7.

For the ith coordinate vector $\boldsymbol{e}_i$, the *partial derivative* is

$$f_{x_i}(\boldsymbol{p}) = \lim_{h \to 0} \frac{f(\boldsymbol{p} + h\boldsymbol{e}_i) - f(\boldsymbol{p})}{h},$$

if the limit exists. Only x_i changes; all other inputs stay fixed. For example, $f_x(a, b)$ differentiates $f(x, b)$ at $x = a$. A *level set* $\{\boldsymbol{x} \in U : f(\boldsymbol{x}) = c\}$ consists of inputs producing the same output. In two dimensions its regular portions are contour curves in input space.

Example 5.1. Choosing which calibration parameter to adjust

The current calibration is $(a, b) = (2, 1)$, with predictions $(1, 2, 3)^\mathsf{T}$ and loss $J = 1$. We want to reduce the mismatch to the fixed observations $(1, 2, 4)^\mathsf{T}$. Expanding (1) gives

$$J = 3a^2 + 2b^2 - 14a - 6b + 21, \qquad J_a = 6a - 14, \quad J_b = 4b - 6.$$

Both sensitivities equal -2 at the current parameters. An offset increase raises all three predictions; a slope increase lowers the first and raises the third, leaving the middle unchanged. Despite these different effects on predictions, their initial loss reductions agree.

For a finite numerical step ε, the exact losses are

$$J(2 + \varepsilon, 1) = 1 - 2\varepsilon + 3\varepsilon^2, \qquad J(2, 1 + \varepsilon) = 1 - 2\varepsilon + 2\varepsilon^2.$$

At $\varepsilon = 0.1$ they are 0.83 and 0.82. The derivative ranks infinitesimal changes; the quadratic terms distinguish these actual updates. Neither negative derivative justifies an arbitrarily large step.

5.2 The gradient and the total derivative

The function f is *differentiable* at $\boldsymbol{p} \in U$ if there is a linear map $L : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(\boldsymbol{p} + \boldsymbol{h}) = f(\boldsymbol{p}) + L(\boldsymbol{h}) + r(\boldsymbol{h}), \qquad \frac{r(\boldsymbol{h})}{\|\boldsymbol{h}\|} \longrightarrow 0 \quad (\boldsymbol{h} \to \boldsymbol{0}). \tag{5.1}$$

We write $r(\boldsymbol{h}) = o(\|\boldsymbol{h}\|)$ and call $L = Df(\boldsymbol{p})$ the *total derivative*.

Reasoning: the linear map determines the gradient. Put $\boldsymbol{h} = t\boldsymbol{e}_i$ in (5.1) and divide the increment by t. The limit gives $f_{x_i}(\boldsymbol{p}) = L(\boldsymbol{e}_i)$. A linear map is determined by its values on a basis, so the derivative is unique and

$$Df(\boldsymbol{p})[\boldsymbol{h}] = \sum_i f_{x_i}(\boldsymbol{p})h_i = \nabla f(\boldsymbol{p})^{\mathsf{T}}\boldsymbol{h}, \qquad \nabla f(\boldsymbol{p}) = \begin{pmatrix} f_{x_1}(\boldsymbol{p}) \\ \vdots \\ f_{x_n}(\boldsymbol{p}) \end{pmatrix}.$$

□

Thus $f(\boldsymbol{p}) + Df(\boldsymbol{p})[\boldsymbol{h}]$ is an affine approximation, while $Df(\boldsymbol{p})$ itself is linear.

A usable sufficient condition

If the first partial derivatives exist and are continuous on a neighbourhood of $\boldsymbol{p}$, then f is differentiable at $\boldsymbol{p}$. We write $f \in C^1(U)$ when these partial derivatives are continuous on U.

Partial derivatives at a single point are insufficient. Set $g(x, y) = xy/(x^2 + y^2)$ away from $(0, 0)$ and $g(0, 0) = 0$. Both coordinate-axis restrictions vanish, so both partial derivatives at the origin are zero. But $g(t, t) = 1/2$ for $t \neq 0$: the function is not even continuous there.

Example 5.2. Sensitivity of a transmitted amplitude

A sensor pipeline records a signal after attenuation. Its model is

$$f(x, y) = xe^{-y}, \qquad x > 0, \quad y \geq 0.$$

Here x is the input amplitude divided by a fixed reference amplitude; y is dimensionless attenuation. The output uses the same amplitude reference. We want to predict how changing either input affects the recorded value near the interior operating point $\boldsymbol{p} = (2, \ln 2)^{\mathsf{T}}$, where $f(\boldsymbol{p}) = 1$.

Differentiation gives $f_x = e^{-y}$, $f_y = -xe^{-y}$ and hence

$$\nabla f(\boldsymbol{p}) = \begin{pmatrix} 1/2 \\ -1 \end{pmatrix}, \qquad \Delta f_{\text{lin}} = \tfrac{1}{2}h_x - h_y.$$

An amplitude increase of 0.01 predicts an output increase of 0.005; an attenuation increase of 0.01 predicts a decrease of 0.01. To compensate for a small attenuation increase h_y, choose $h_x = 2h_y$ to first order. This is a local control rule; finite compensation will require curvature.

5.3 Directions and the steepest local change

For a unit vector $\boldsymbol{d}$, define

$$D_{\boldsymbol{d}}f(\boldsymbol{p}) = \frac{\mathrm{d}}{\mathrm{d}s} f(\boldsymbol{p} + s\boldsymbol{d})\bigg|_{s=0}.$$

If f is differentiable, substitution into (5.1) gives

$$D_{\boldsymbol{d}}f(\boldsymbol{p}) = \nabla f(\boldsymbol{p})^{\mathsf{T}}\boldsymbol{d}. \tag{5.2}$$

Since $\|\boldsymbol{d}\| = 1$, this is a rate per unit distance in the chosen coordinates. A nonunit vector also specifies a speed and gives a rate per unit of its path parameter.

Reasoning: steepest change follows from Cauchy–Schwarz. For $\|\boldsymbol{d}\| = 1$,

$$-\|\nabla f(\boldsymbol{p})\| \le \nabla f(\boldsymbol{p})^{\mathsf{T}}\boldsymbol{d} \le \|\nabla f(\boldsymbol{p})\|.$$

If $\nabla f(\boldsymbol{p}) \neq \boldsymbol{0}$, the bounds are attained at $\boldsymbol{d} = \mp\nabla f(\boldsymbol{p})/\|\nabla f(\boldsymbol{p})\|$, respectively. The negative gradient gives steepest first-order decrease. If the gradient vanishes, all first-order rates vanish; the function need not be constant nearby. □

At $(2,1)$, $\nabla J = (-2,-2)^{\mathsf{T}}$. The unit direction $(1,1)^{\mathsf{T}}/\sqrt{2}$ has rate $-2\sqrt{2}$; the perpendicular direction $(1,-1)^{\mathsf{T}}/\sqrt{2}$ has rate zero. A tangent direction gives no first-order change, but a finite straight step may leave the contour.

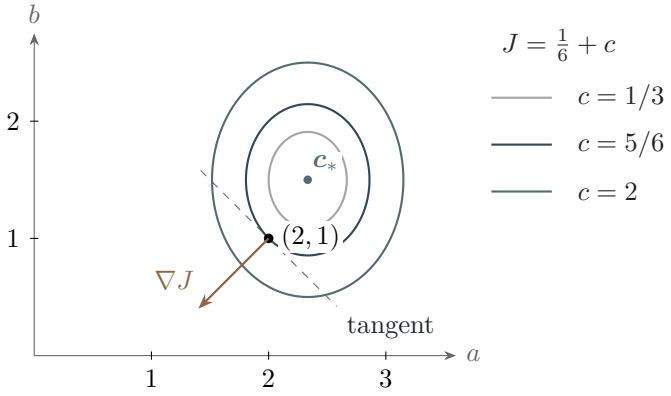


FIGURE 5.1. Exact contours $J = 1/6 + c$ for $c = 1/3, 5/6, 2$. The middle contour passes through $(2,1)$. Its tangent has direction $(1,-1)^{\mathsf{T}}$; the gradient points towards increasing loss and is drawn at 0.3 of its actual vector length. Coordinate scales are equal.

5.4 A moving input and the chain rule

Let $\boldsymbol{x}(s)$ be differentiable at s_0, with $\boldsymbol{x}(s_0) = \boldsymbol{p}$, and let f be differentiable at $\boldsymbol{p}$. Then

$$\frac{\mathrm{d}}{\mathrm{d}s} f(\boldsymbol{x}(s))\bigg|_{s=s_0} = \nabla f(\boldsymbol{p})^{\mathsf{T}}\boldsymbol{x}'(s_0). \tag{5.3}$$

Each sensitivity is multiplied by the corresponding input velocity.

Reasoning: compose the local linear maps. The input displacement is $\boldsymbol{x}(s_0 + \delta s) - \boldsymbol{p} = \boldsymbol{x}'(s_0)\delta s + o(|\delta s|)$. Insert it into (5.1). The output increment becomes $Df(\boldsymbol{p})[\boldsymbol{x}'(s_0)]\delta s + o(|\delta s|)$. Division by δs gives (5.3). □

The same definition applies to a vector output $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$:

$$F(\boldsymbol{p} + \boldsymbol{h}) = F(\boldsymbol{p}) + DF(\boldsymbol{p})\boldsymbol{h} + o(\|\boldsymbol{h}\|).$$

Here the remainder is small in Euclidean norm and the derivative is the $m \times n$ *Jacobian matrix*

$$DF(\boldsymbol{p}) = \left(\frac{\partial F_i}{\partial x_j}(\boldsymbol{p})\right)_{i,j}.$$

For differentiable F at $\boldsymbol{p}$ and G at $F(\boldsymbol{p})$, the same composition argument gives $D(G \circ F)(\boldsymbol{p}) = DG(F(\boldsymbol{p}))\, DF(\boldsymbol{p})$. Rows correspond to outputs, columns to inputs.

In particular, if $\boldsymbol{x} = S\boldsymbol{z}$ for an invertible diagonal scale matrix, then $\nabla_{\boldsymbol{z}}(f \circ S) = S^{\mathsf{T}}\nabla_{\boldsymbol{x}} f(S\boldsymbol{z})$. Changing reference units changes the Euclidean notion of a unit step, and can change the physical direction represented by steepest descent.

Example 5.3. A control path has both direction and speed

In the attenuation model, suppose a control parameter s increases amplitude and attenuation together:

$$\boldsymbol{x}(s) = (2 + 3s, \ln 2 + 4s)^{\mathsf{T}} \quad \text{for sufficiently small } s.$$

The velocity is $(3, 4)^{\mathsf{T}}$, of length 5. At $s = 0$,

$$\frac{\mathrm{d}}{\mathrm{d}s} f(\boldsymbol{x}(s)) = (1/2, -1) \begin{pmatrix} 3 \\ 4 \end{pmatrix} = -\frac{5}{2}.$$

The amplitude contribution is $3/2$, the attenuation contribution -4; their sum is negative. Per unit input distance, the rate is instead $D_{\boldsymbol{d}} f = -1/2$ for $\boldsymbol{d} = (3/5, 4/5)^{\mathsf{T}}$. Direct substitution checks both interpretations: $f(\boldsymbol{x}(s)) = (1+\frac{3}{2}s)e^{-4s}$. The control increases the input amplitude but initially decreases the recorded output.

If a differentiable path stays on a level set, $f(\boldsymbol{x}(s)) = c$, then the chain rule gives $\nabla f^{\mathsf{T}} \boldsymbol{x}' = 0$. At points with nonzero gradient, the gradient is normal to every tangent velocity. This will be the geometric basis of multipliers.

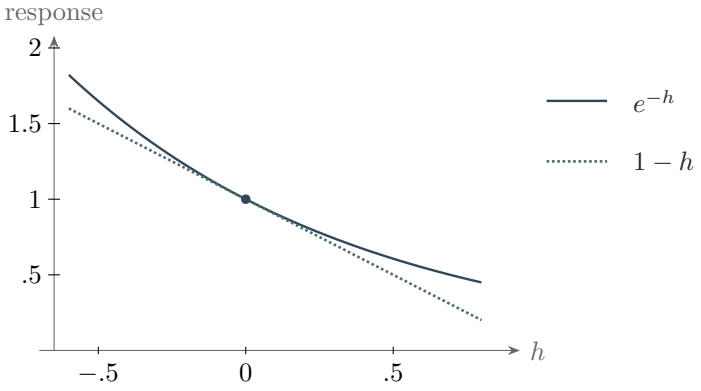


FIGURE 5.2. With amplitude fixed at $x = 2$, an attenuation change h gives the exact output e^{-h}. The tangent approximation $1 - h$ agrees in value and derivative at zero. Its visible separation from the curve records the accumulated effect of curvature.

5.5 Finite changes and uncertainty

Differentiability gives the first-order prediction $\Delta f_{\text{lin}} = \nabla f(\boldsymbol{p})^{\mathsf{T}} \boldsymbol{h}$. Its error is $o(\|\boldsymbol{h}\|)$, an asymptotic statement rather than a numerical bound at a specified step.

Example 5.4. Predicting an update and checking a tolerance

For the same sensor, take $\boldsymbol{p} = (2, \ln 2)^{\mathsf{T}}$ and the known update $\boldsymbol{h} = (0.04, 0.01)^{\mathsf{T}}$. The predicted output is

$$f(\boldsymbol{p}) + \Delta f_{\text{lin}} = 1 + \tfrac{1}{2}(0.04) - 0.01 = 1.01.$$

The exact output is $2.04e^{-\ln 2 - 0.01} = 1.02e^{-0.01} \approx 1.0098508304$. The approximation overestimates it by about $1.49 \cdot 10^{-4}$.

Now suppose the changes are unknown, with only $|h_x| \le 0.04$ and $|h_y| \le 0.01$ guaranteed. The worst-case linearized deviation is

$$|\Delta f_{\text{lin}}| \le \tfrac{1}{2}|h_x| + |h_y| \le 0.03.$$

This does not certify a 3% output tolerance. Since f increases with x and decreases with y on the error rectangle, the exact interval is

$$0.98e^{-0.01} \le f(\boldsymbol{p}+\boldsymbol{h}) \le 1.02e^{0.01}.$$

The upper endpoint is about 1.03025, slightly outside $[0.97, 1.03]$. The known update and the unknown error bounds answer different questions; neither uses a probability model.

For a Euclidean error budget $\|\boldsymbol{h}\| \le \varepsilon$, Cauchy–Schwarz instead gives $|\Delta f_{\text{lin}}| \le \|\nabla f(\boldsymbol{p})\|\,\varepsilon$. Componentwise bounds describe a box; a norm bound describes a ball.

5.6 Further perspective: amplitude and phase sensitivities

Example 5.5. Choosing samples to distinguish two signal parameters

A periodic sensor output is modelled by $s(A, \phi; t) = A\cos(\omega t - \phi)$. The angular frequency $\omega > 0$ and sample times t are known; amplitude $A > 0$ and phase ϕ are unknown. Phase is measured in radians: a positive ϕ shifts a cosine peak to $t = \phi/\omega$. We want to estimate small parameter changes near $(A, \phi) = (2, 0)$.

At a fixed sample time,

$$\frac{\partial s}{\partial A} = \cos(\omega t - \phi), \qquad \frac{\partial s}{\partial \phi} = A\sin(\omega t - \phi).$$

Take samples at $\omega t = 0$ and $\pi/2$. Their output vector and Jacobian at the operating point are

$$S(A, \phi) = \begin{pmatrix} A\cos\phi \\ A\sin\phi \end{pmatrix}, \qquad DS(2, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Thus observed changes $(\delta s_1, \delta s_2)$ give the linearized estimates $\delta A = \delta s_1$ and $\delta\phi = \delta s_2/2$. For changes $(0.03, 0.02)$, these are 0.03 amplitude units and 0.01 radians.

The peak sample measures amplitude to first order; the zero-crossing sample measures phase. Full rank makes the two parameter changes distinguishable in this linear model. A peak sample alone has zero first-order phase sensitivity, so it cannot supply the same estimate. These estimates are local; phase remains defined modulo 2π.

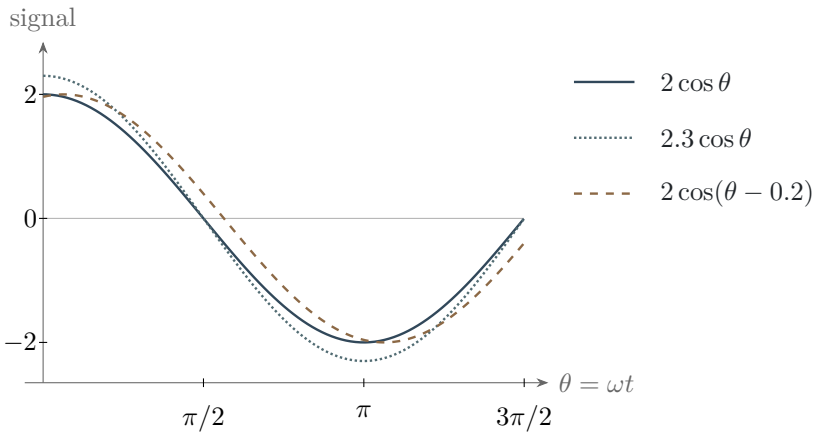


FIGURE 5.3. Exact waveforms on common scales. Amplitude changes the peak size; positive phase shifts the waveform right. Derivatives describe the local limit.

5.7 Questions and exercises

Questions for review

T5.1. Define a partial derivative and identify what is held fixed. In a fitting problem, distinguish measured inputs, observed outputs, and optimized parameters.

T5.2. State differentiability using a linear map and a remainder. Give a practical sufficient condition and explain why partial derivatives at one point alone do not suffice.

T5.3. Derive the directional derivative formula for a differentiable function. Why does a direction need normalization when rates per unit distance are compared?

T5.4. Use Cauchy–Schwarz to identify the steepest local increase and decrease. How does changing coordinate scales affect this statement?

T5.5. Define the Jacobian of a vector-valued map and state the chain rule. Use its path version to show that a level curve's tangent is orthogonal to the gradient.

T5.6. Distinguish an exact finite change, its first-order prediction, and a linearized uncertainty bound. How do box and ball error constraints differ?

Exercises

5.1. For $f(x, y) = x^2y + e^y$, compute both partial derivatives and the gradient at $(1, 0)$. Predict the change caused by $(h_x, h_y) = (0.02, -0.01)$.

5.2. At $(2, \ln 2)$ for $f(x, y) = xe^{-y}$, find the directional derivatives in the directions $(1, 0)^\mathsf{T}$, $(0, 1)^\mathsf{T}$, and $(1, 2)^\mathsf{T}/\sqrt{5}$. Find the greatest first-order increase per unit distance and a direction attaining it.

5.3. Let $f(x, y) = x^2 + 2y^2$ and $\boldsymbol{x}(s) = (1 + s, 1 - s)^\mathsf{T}$. Compute the output rate at $s = 0$ both by the chain rule and by direct substitution. Repeat per unit distance travelled.

5.4. For the calibration loss J, compute $\nabla J(2, 1)$. Find a unit direction with zero initial change. Substitute a finite displacement $h(1, -1)^\mathsf{T}$ and show that the exact loss is not constant along this line.

5.5. For $f(x, y) = xy$ at $(2, 3)$, compare the exact and linearized values after the displacement $(0.1, -0.2)$. Express the exact remainder for an arbitrary displacement (h_x, h_y).

5.6. For $f(x, y) = xe^{-y}$ at $(2, \ln 2)$, obtain the linearized worst-case change under $|h_x| \leq 0.02$, $|h_y| \leq 0.03$. Obtain an exact interval for the output over this rectangle using monotonicity.

5.7. (*) For $s(A, \phi; t) = A\cos(\omega t - \phi)$, find the parameter derivative matrix at $(2, 0)$ for samples $\omega t = 0, \pi$. Compare its rank with the sample pair $0, \pi/2$, and interpret the difference.

5.8. (*) Write $f(x, y) = x^2 + y^2$ in coordinates $x = 10u$, $y = v$. At the physical point $(x, y) = (1, 1)$, compute the negative gradient in each coordinate system. Map the (u, v) direction back to (x, y) and explain why the two resulting directions differ.

Lecture 6

Curvature, Taylor models, and unconstrained optima

The next term after a linear derivative is a quadratic form. Its mixed terms describe interacting sensitivities; its signs distinguish an optimum from a saddle.

6.1 The Hessian and the second-order model

We write $f \in C^2(U)$ when all partial derivatives through order two are continuous on the open set U. Its *Hessian* is

$$H_f(\boldsymbol{x}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}) \right)_{i,j}.$$

Continuity of the second partial derivatives ensures equality of the mixed partials, hence symmetry of H_f. In two variables,

$$\tfrac{1}{2}\boldsymbol{h}^{\mathsf{T}} H_f(\boldsymbol{p})\boldsymbol{h} = \tfrac{1}{2} f_{xx} h_x^2 + f_{xy} h_x h_y + \tfrac{1}{2} f_{yy} h_y^2,$$

with derivatives evaluated at $\boldsymbol{p}$. The mixed term records how one input changes the sensitivity to the other.

Taylor's formula to second order

If $f \in C^2$ on a neighbourhood of $\boldsymbol{p}$, then

$$f(\boldsymbol{p}+\boldsymbol{h}) = f(\boldsymbol{p}) + \nabla f(\boldsymbol{p})^{\mathsf{T}}\boldsymbol{h} + \tfrac{1}{2}\boldsymbol{h}^{\mathsf{T}} H_f(\boldsymbol{p})\boldsymbol{h} + o(\|\boldsymbol{h}\|^2). \tag{6.1}$$

The remainder divided by $\|\boldsymbol{h}\|^2$ tends to zero as $\boldsymbol{h} \to \boldsymbol{0}$. For a quadratic function the formula is exact without the remainder.

Reasoning: apply the one-variable theorem on a line. For a sufficiently small $\boldsymbol{h}$, put $q(s) = f(\boldsymbol{p} + s\boldsymbol{h})$, $0 \le s \le 1$. The chain rule gives

$$q'(s) = \nabla f(\boldsymbol{p}+s\boldsymbol{h})^{\mathsf{T}}\boldsymbol{h}, \qquad q''(s) = \boldsymbol{h}^{\mathsf{T}} H_f(\boldsymbol{p}+s\boldsymbol{h})\boldsymbol{h}.$$

One-variable Taylor gives $q(1) = q(0) + q'(0) + q''(\xi)/2$ for some $0 < \xi < 1$. By continuity, every entry of $H_f(\boldsymbol{p}+\xi\boldsymbol{h}) - H_f(\boldsymbol{p})$ tends to zero as $\boldsymbol{h} \to \boldsymbol{0}$, uniformly along this shrinking segment. Its quadratic contribution is $o(\|\boldsymbol{h}\|^2)$. □

This argument also gives a finite error bound. If $|\boldsymbol{v}^{\mathsf{T}} H_f(\boldsymbol{x})\boldsymbol{v}| \le M\|\boldsymbol{v}\|^2$ for every $\boldsymbol{v}$ and every $\boldsymbol{x}$ on the segment, then

$$\left| f(\boldsymbol{p}+\boldsymbol{h}) - f(\boldsymbol{p}) - \nabla f(\boldsymbol{p})^{\mathsf{T}}\boldsymbol{h} \right| \le \tfrac{1}{2} M \|\boldsymbol{h}\|^2.$$

By Part I, it suffices to bound the absolute values of all Hessian eigenvalues along that segment by M.

Example 6.1. How accurate is amplitude compensation?

For $f(x,y) = xe^{-y}$ at $\boldsymbol{p} = (2, \ln 2)^{\mathsf{T}}$, differentiation gives

$$H_f(x,y) = \begin{pmatrix} 0 & -e^{-y} \\ -e^{-y} & xe^{-y} \end{pmatrix}, \qquad H_f(\boldsymbol{p}) = \begin{pmatrix} 0 & -1/2 \\ -1/2 & 1 \end{pmatrix}.$$

The second-order prediction is therefore

$$f(\boldsymbol{p}+\boldsymbol{h}) \approx 1 + \tfrac{1}{2}h_x - h_y - \tfrac{1}{2}h_x h_y + \tfrac{1}{2}h_y^2.$$

For the known update $(0.04, 0.01)$ of Example 5.4, the linear change is 0.01 and the quadratic correction is -0.00015:

first-order output:	1.01,
second-order output:	1.00985,
exact output:	$1.02e^{-0.01} \approx 1.0098508304$.

The mixed term is negative because increased attenuation weakens the benefit of increased amplitude.

Now use the first-order compensation rule $h_x = 2h_y$, taking $(h_x, h_y) = (0.04, 0.02)$. The linear change vanishes, but the quadratic change is -0.0002. The exact output is $1.02e^{-0.02} \approx 0.9998026468$. To maintain output exactly at 1, the amplitude must instead be $x = 2e^{h_y}$. Thus the rule $h_x = 2h_y$ omits a second-order adjustment.

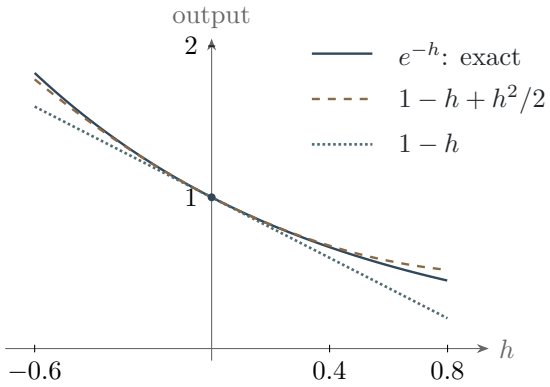


FIGURE 6.1. At fixed $x = 2$, changing y from $\ln 2$ to $\ln 2 + h$ gives the exact output e^{-h}. The quadratic model includes the curvature omitted by the tangent line. The claim of improved approximation is local near $h = 0$.

6.2 Stationary points and necessary conditions

A point $\boldsymbol{p}$ is a *local minimum* on a domain D if some $\varepsilon > 0$ satisfies $f(\boldsymbol{x}) \geq f(\boldsymbol{p})$ for all $\boldsymbol{x} \in D$ with $\|\boldsymbol{x} - \boldsymbol{p}\| < \varepsilon$. It is *strict* if the inequality is strict for $\boldsymbol{x} \neq \boldsymbol{p}$ there, and *global* if the inequality holds on all of D. Reverse inequalities for maxima. Local need not mean global: $u^3 - 3u$ has a local minimum at $u = 1$ but tends to $-\infty$ as $u \to -\infty$.

At a differentiable *interior* local extremum,

$$\nabla f(\boldsymbol{p}) = \boldsymbol{0}. \tag{6.2}$$

A point satisfying this equation is called *stationary*.

Reasoning: restrict to both signs of each coordinate change. For $q(t) = f(\boldsymbol{p} + t\boldsymbol{e}_i)$, small positive and negative t are admissible because $\boldsymbol{p}$ is interior. A local extremum of f gives one of q at zero. The one-variable necessary condition yields $q'(0) = f_{x_i}(\boldsymbol{p}) = 0$ for each i. □

For $f \in C^2$, a local minimum also requires $\boldsymbol{v}^\mathsf{T} H_f(\boldsymbol{p})\boldsymbol{v} \geq 0$ for every $\boldsymbol{v}$: apply the one-variable second-derivative necessary condition to $f(\boldsymbol{p} + t\boldsymbol{v})$. Thus $H_f(\boldsymbol{p})$ must be positive semidefinite; at a maximum it must be negative semidefinite. A *saddle point* here means a stationary point with both larger and smaller values arbitrarily nearby.

6.3 What the Hessian can decide

At a stationary point Taylor's linear term vanishes. For a real symmetric Hessian, Part I gives an orthonormal eigenbasis and

$$H_f(\boldsymbol{p}) = Q\Lambda Q^\mathsf{T}, \qquad \boldsymbol{h}^\mathsf{T} H_f(\boldsymbol{p})\boldsymbol{h} = \sum_i \lambda_i z_i^2, \quad \boldsymbol{z} = Q^\mathsf{T}\boldsymbol{h}.$$

Hence the signs of the eigenvalues describe curvature in all directions, not merely along the coordinate axes.

Second-order classification

Assume $f \in C^2$ near an interior stationary point $\boldsymbol{p}$.

- All $\lambda_i > 0$ (positive definite Hessian): strict local minimum.
- All $\lambda_i < 0$ (negative definite Hessian): strict local maximum.
- Both a positive and a negative λ_i (indefinite Hessian): saddle.
- A semidefinite Hessian with a zero eigenvalue: the test is inconclusive.

An indefinite Hessian still decides a saddle if some additional eigenvalues are zero.

Reasoning: positive curvature dominates the remainder. If $\mu = \min_i \lambda_i > 0$, then $\boldsymbol{h}^\mathsf{T} H_f(\boldsymbol{p})\boldsymbol{h} \geq \mu \|\boldsymbol{h}\|^2$. For sufficiently small nonzero $\boldsymbol{h}$, the Taylor remainder has absolute value at most $\mu \|\boldsymbol{h}\|^2 / 4$. Consequently $f(\boldsymbol{p} + \boldsymbol{h}) - f(\boldsymbol{p}) \geq \mu \|\boldsymbol{h}\|^2 / 4 > 0$. Apply this to $-f$ for a maximum. For an indefinite Hessian, short steps along a positive and a negative eigenvector give opposite signs of change. □

For a 2×2 Hessian, let $D = f_{xx} f_{yy} - f_{xy}^2$, evaluated at the stationary point. If $D > 0$, $f_{xx} > 0$ gives a minimum and $f_{xx} < 0$ a maximum. If $D < 0$, the point is a saddle; if $D = 0$, this shortcut is inconclusive. A determinant alone does not classify a Hessian in arbitrary dimension.

Even two positive diagonal entries do not suffice: $f = x^2 + 4xy + y^2$ has Hessian $\left(\begin{smallmatrix} 2 & 4 \\ 4 & 2 \end{smallmatrix}\right)$ with eigenvalues $6, -2$. Indeed $f(t, t) = 6t^2$ but $f(t, -t) = -2t^2$.

Example 6.2. A stationary fit that fails to minimize the loss

A calibration procedure measures squared gain and a separate offset. In fixed reference units, its model predicts (x^2, y); the target readings are $(1, 0)$. The signed gain x and offset y are unknown. With equal residual weights, the fitting loss is

$$f(x, y) = (x^2 - 1)^2 + y^2, \qquad (x, y) \in \mathbb{R}^2.$$

The gradient and Hessian are

$$\nabla f = \begin{pmatrix} 4x(x^2 - 1) \\ 2y \end{pmatrix}, \qquad H_f = \operatorname{diag}(12x^2 - 4, 2).$$

Stationarity gives $(1,0)$, $(-1,0)$, and $(0,0)$. At the first two points the Hessian is $\operatorname{diag}(8,2)$, so each is a strict local minimum. Both are global: the loss is nonnegative and equals zero there. The data identify the magnitude of the gain but not its sign; a sign-sensitive measurement would be needed to distinguish the fits.

At $(0,0)$, $H_f = \operatorname{diag}(-4,2)$ and the loss is 1. This is a saddle:

$$f(h,0) - 1 = -2h^2 + h^4, \qquad f(0,h) - 1 = h^2.$$

A zero gradient at zero gain is therefore no certificate of a good fit. Unlike linear least squares, a squared-residual loss for a nonlinear prediction map need not be convex.

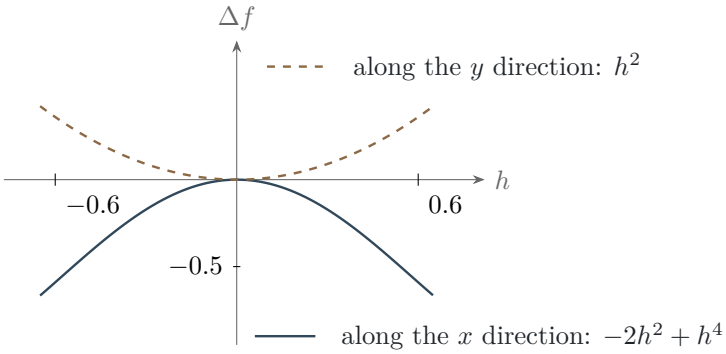


FIGURE 6.2. Two exact slices through the stationary fit $(0,0)$ for $(x^2-1)^2 + y^2$. The vertical axis shows change from its loss 1. Opposite signs of change demonstrate the saddle.

Example 6.3. Flat curvature can conceal a unique minimum

If two energy readings are both zero and their predicted values are x^2, y^2, their squared-residual loss is $f_1 = x^4 + y^4$. At the origin the gradient and Hessian vanish, yet the origin is the unique global minimizer: both fourth powers are nonnegative and their sum vanishes only there.

To see why the zero Hessian alone cannot decide, compare $f_2 = -x^4 - y^4$ and $f_3 = x^4 - y^4$. They have the same zero gradient and Hessian at the origin, but give a strict global maximum and a saddle, respectively.

6.4 Least squares: curvature from the design matrix

Let $A \in \mathbb{R}^{m\times n}$ and $\boldsymbol{y} \in \mathbb{R}^m$ be fixed, with $\boldsymbol{r}(\boldsymbol{c}) = \boldsymbol{y} - A\boldsymbol{c}$ and $J(\boldsymbol{c}) = \|\boldsymbol{r}(\boldsymbol{c})\|^2$.

Reasoning: read the derivatives from an exact expansion. For any parameter displacement $\boldsymbol{h}$,

$$\begin{aligned} J(\boldsymbol{c}+\boldsymbol{h}) &= \|\boldsymbol{r}(\boldsymbol{c}) - A\boldsymbol{h}\|^2 \\ &= J(\boldsymbol{c}) - 2\boldsymbol{r}(\boldsymbol{c})^{\mathsf{T}} A\boldsymbol{h} + \boldsymbol{h}^{\mathsf{T}} A^{\mathsf{T}} A\boldsymbol{h}. \end{aligned}$$

Comparison with Taylor's formula gives

$$\nabla J(\boldsymbol{c}) = -2A^{\mathsf{T}}\boldsymbol{r}(\boldsymbol{c}), \qquad H_J = 2A^{\mathsf{T}}A. \tag{6.3}$$

Thus stationarity is exactly the normal equation $A^{\mathsf{T}}A\boldsymbol{c} = A^{\mathsf{T}}\boldsymbol{y}$, or residual orthogonality $A^{\mathsf{T}}\boldsymbol{r} = \boldsymbol{0}$. □

The Gram matrix $G = A^\mathsf{T}A$ satisfies $\boldsymbol{h}^\mathsf{T}G\boldsymbol{h} = \|A\boldsymbol{h}\|^2$ and $\ker G = \ker A$. It is positive definite precisely when the columns of A are independent. At any normal-equation solution $\boldsymbol{c}_*$, the expansion reduces to

$$J(\boldsymbol{c}_* + \boldsymbol{h}) = J(\boldsymbol{c}_*) + \|A\boldsymbol{h}\|^2 .$$

This proves global minimality directly. The fitted prediction is unique, as in Part I; the coefficient vector is unique exactly when $\ker A = \{\mathbf{0}\}$. Otherwise displacements in $\ker A$ leave every prediction unchanged.

Example 6.4. Parameter drift and a calibration loss budget

For the temperature calibration,

$$\nabla J = \begin{pmatrix} 6a - 14 \\ 4b - 6 \end{pmatrix}, \qquad H_J = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}.$$

The eigenvalues $6, 4$ are positive. Stationarity recovers $(a_*, b_*) = (7/3, 3/2)$, and completion of squares gives the exact identity

$$J(a, b) = \tfrac{1}{6} + 3(a - 7/3)^2 + 2(b - 3/2)^2.$$

Suppose recalibration is required once parameter drift adds more than a loss budget $\rho > 0$. The acceptable drift $\boldsymbol{h}$ obeys $3h_a^2 + 2h_b^2 \le \rho$, an ellipse centred at the optimum. An offset-only drift may have magnitude at most $\sqrt{\rho/3}$; a slope-only drift at most $\sqrt{\rho/2}$. The larger offset curvature means a tighter allowance in these fixed coordinates. The residual floor $1/6$ is separate from this additional loss.

6.5 Convexity: when a local optimum is global

A set U is *convex* if it contains every segment joining two of its points. A function $f : U \to \mathbb{R}$ is *convex* if

$$f((1-t)\boldsymbol{x} + t\boldsymbol{z}) \le (1-t)f(\boldsymbol{x}) + tf(\boldsymbol{z}) \quad (\boldsymbol{x}, \boldsymbol{z} \in U,\ 0 \le t \le 1).$$

It is *strictly convex* if the inequality is strict for $\boldsymbol{x} \ne \boldsymbol{z}$ and $0 < t < 1$. For $f \in C^2$ on an open convex U, convexity is equivalent to $H_f(\boldsymbol{x})$ being positive semidefinite everywhere on U. Indeed, each segment restriction has second derivative $\boldsymbol{v}^\mathsf{T}H_f\boldsymbol{v}$, so the criterion reduces to the one-variable convexity criterion. Positive definiteness everywhere implies strict convexity.

Reasoning: a tangent plane bounds a differentiable convex function. For $0 < t < 1$, convexity gives

$$\frac{f(\boldsymbol{x} + t(\boldsymbol{z} - \boldsymbol{x})) - f(\boldsymbol{x})}{t} \le f(\boldsymbol{z}) - f(\boldsymbol{x}).$$

Let $t \to 0$. Differentiability yields $f(\boldsymbol{z}) \ge f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^\mathsf{T}(\boldsymbol{z} - \boldsymbol{x})$. At a stationary point the last term is zero, proving global minimality. □

A strictly convex function has at most one minimizer: two distinct minimizers would have a midpoint of smaller value. Existence is a separate question; e^x is strictly convex on $\mathbb{R}$ but has no minimizer. Linear least squares always has a best prediction by the projection argument, with unique parameters when the design columns are independent.

6.6 Further perspective: choosing an optimization step

Example 6.5. A descent direction still needs a step length

Start calibration at $\boldsymbol{c}_0 = (2,1)^\mathsf{T}$, where $J = 1$ and $\nabla J = (-2,-2)^\mathsf{T}$. A gradient update uses

$$\boldsymbol{c}(\eta) = \boldsymbol{c}_0 - \eta \nabla J(\boldsymbol{c}_0) = (2+2\eta, 1+2\eta)^\mathsf{T}, \qquad \eta > 0.$$

The parameter η scales the gradient; the actual displacement length is $2\sqrt{2}\,\eta$. Substitution gives

$$J(\boldsymbol{c}(\eta)) = 1 - 8\eta + 20\eta^2.$$

The update decreases loss exactly when $0 < \eta < 2/5$. At $\eta = 1/4$, the loss is $1/4$; at $\eta = 1/2$, it rises to 2. The best point on this line has $\eta = 1/5$ and loss $1/5$, still above the unrestricted minimum $1/6$.

A Newton step instead minimizes the local quadratic model by solving

$$H_J \boldsymbol{s} = -\nabla J(\boldsymbol{c}_0), \qquad \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix} \boldsymbol{s} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

It gives $\boldsymbol{s} = (1/3, 1/2)^\mathsf{T}$ and reaches $\boldsymbol{c}_*$ in one step because this strictly convex objective is exactly quadratic. For a nonlinear objective, the Taylor model is local; an invertible Hessian alone does not ensure that its stationary point is a minimum or that a full step decreases the true loss.

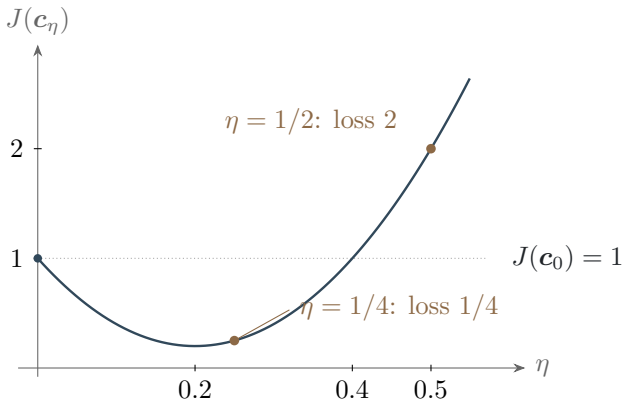


FIGURE 6.3. Exact loss after a first gradient step from $(2,1)$. The initial slope is negative, but a sufficiently large positive step raises the loss. The horizontal dotted line marks the starting value.

6.7 Questions and exercises

Questions for review

T6.1. Define the Hessian and state a sufficient condition for its symmetry. Explain the coefficient of the mixed term in the quadratic Taylor model.

T6.2. State the second-order Taylor formula with its assumptions and remainder. Explain the reduction to a one-variable function along a line.

T6.3. Distinguish local, strict local, and global minima. Why is a zero gradient necessary only for a differentiable interior extremum?

T6.4. State the necessary and sufficient second-order conditions used here. Why is a semidefinite Hessian with zero eigenvalues inconclusive?

T6.5. Derive the gradient and Hessian of a least-squares objective. Relate positive definiteness to the kernel of the design matrix and to uniqueness of fitted parameters.

T6.6. Distinguish a local curvature test from a convexity argument on the whole domain. Can a convex function fail to attain its infimum?

Exercises

6.1. Find the second-order Taylor polynomial of $f(x, y) = e^x + y^2$ at $(0, 1)$. Use it to estimate $f(0.1, 0.9)$ and compare with the exact expression.

6.2. For $f(x, y) = x^2 + 4xy + y^2$, compute the Hessian, its eigenvalues, and two unit eigenvectors. Classify the origin and calculate exact changes along those directions.

6.3. Find and classify all stationary points of $f(x, y) = x^3 - 3x + 2y^2$. Is any of them a global minimum on $\mathbb{R}^2$? Justify your answer from function values, not only from the Hessian.

6.4. Compare $f_+(x, y) = x^4 + y^2$ and $f_-(x, y) = -x^4 + y^2$ at the origin. They have the same Hessian there. Classify each point directly and identify the direction in which the quadratic model is silent.

6.5. For $J(a, b) = (1 - a + b)^2 + (2 - a)^2 + (4 - a - b)^2$, differentiate the sum and verify $\nabla J = -2A^\mathsf{T}\boldsymbol{r}$, where $\boldsymbol{r} = \boldsymbol{y} - A\boldsymbol{c}$. Find the minimizer and prove its global uniqueness by completing squares.

6.6. Suppose $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ and $\boldsymbol{y} = (1, 2)^\mathsf{T}$. Find all least-squares minimizers, the Hessian, and its kernel. Explain why the loss is convex but the minimizing parameters are not unique.

6.7. (*) In Example 6.5, minimize the loss over all $\eta \geq 0$ along the first gradient search line. Compare this minimum with the unconstrained minimum $1/6$ and explain the gap geometrically.

6.8. (*) For $F(x, y) = \frac{1}{2}(x^2 + 4y^2) - x - 4y$, start at $(0, 0)$. Compute one gradient step with $\eta = 1/4$ and one Newton step. Compare their losses with the true minimum and explain why the Newton step is exact here.

Lecture 7

Equality-constrained optimization

An exact requirement restricts the inputs that may be compared. A constrained optimum need only resist changes that respect that requirement.

7.1 Feasible points and tangent directions

For $g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ and a fixed c, the *feasible set* is $M = \{\boldsymbol{x} \in U : g(\boldsymbol{x}) = c\}$. A constrained local minimum satisfies the local-minimum inequality only for points of M. Initially U is open and there are no additional active inequality restrictions.

Assume $g \in C^1(U)$. If a differentiable feasible path satisfies $\boldsymbol{x}(0) = \boldsymbol{p}$, differentiation of $g(\boldsymbol{x}(s)) = c$ gives $\nabla g(\boldsymbol{p})^\mathsf{T} \boldsymbol{x}'(0) = 0$. The point is *regular* if $\nabla g(\boldsymbol{p}) \neq \boldsymbol{0}$. At a regular point the level set is locally a smooth surface, with tangent space

$$T_{\boldsymbol{p}} = \ker(\nabla g(\boldsymbol{p})^\mathsf{T}) = \{\boldsymbol{v} : \nabla g(\boldsymbol{p})^\mathsf{T} \boldsymbol{v} = 0\}. \tag{7.1}$$

Every vector in this space is the velocity of a feasible curve. We use this regular-level-set fact without proving the implicit function theorem.

For $x + y = 1$, the normal is $(1, 1)^\mathsf{T}$ and the tangent direction is $(1, -1)^\mathsf{T}$; a tangent step stays on the line. On the unit circle at $(1, 0)$, $(0, 1)^\mathsf{T}$ is tangent, but $(1, h)$ is not feasible for $h \neq 0$. A tangent vector describes first-order motion, not an exact straight path on a curved set.

7.2 Why multipliers appear

Reasoning: use the orthogonal complement of the tangent space. Let $f, g \in C^1(U)$ and let $\boldsymbol{p}$ be a regular constrained local extremum. The derivative of f along every feasible curve through $\boldsymbol{p}$ vanishes. Hence $\nabla f(\boldsymbol{p})^\mathsf{T} \boldsymbol{v} = 0$ for all $\boldsymbol{v} \in T_{\boldsymbol{p}}$. Part I gives $(\ker \boldsymbol{a}^\mathsf{T})^\perp = \operatorname{span}\{\boldsymbol{a}\}$ for $\boldsymbol{a} \neq \boldsymbol{0}$. Apply this with $\boldsymbol{a} = \nabla g(\boldsymbol{p})$ to obtain

$$\nabla f(\boldsymbol{p}) = \lambda \nabla g(\boldsymbol{p}), \qquad g(\boldsymbol{p}) = c. \tag{7.2}$$

The scalar λ is the *Lagrange multiplier*. □

Equations for candidates

With the fixed sign convention

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda\big(g(\boldsymbol{x}) - c\big),$$

solve $\nabla_{\boldsymbol{x}} L = \boldsymbol{0}$ and $g(\boldsymbol{x}) = c$. These equations are necessary at regular extrema under the assumptions above.

Prove optimality on the feasible set, for example by substitution. Active inequality boundaries and singular constraints need separate treatment. The candidate need not minimize L in $(\boldsymbol{x}, \lambda)$!

Example 7.1. Reconciling two estimates with an exact total

Two channels report normalized signed deviations $(2, 1)$, while an independent balance fixes their true sum at 1. We seek corrected estimates (x, y) closest to these reports. The channels use the same reference scale and equal squared-error weights:

$$F(x, y) = \tfrac{1}{2}((x-2)^2 + (y-1)^2), \qquad x + y = 1.$$

The factor $1/2$ simplifies derivatives and does not change the minimizer. Since $\nabla g = (1, 1)^\mathsf{T} \neq \mathbf{0}$, the multiplier equations are

$$x - 2 = \lambda, \qquad y - 1 = \lambda, \qquad x + y = 1.$$

Thus $3 + 2\lambda = 1$, giving $\lambda = -1$ and $(x, y) = (1, 0)$ with $F = 1$. Every feasible pair is $(1 + s, -s)$, and

$$F(1 + s, -s) = 1 + s^2.$$

This proves unique global optimality. Equal weights distribute the total discrepancy -2 equally between the channels. The correction $(-1, -1)^\mathsf{T}$ is perpendicular to $(1, -1)^\mathsf{T}$: this is projection onto an *affine* line, a shifted subspace.

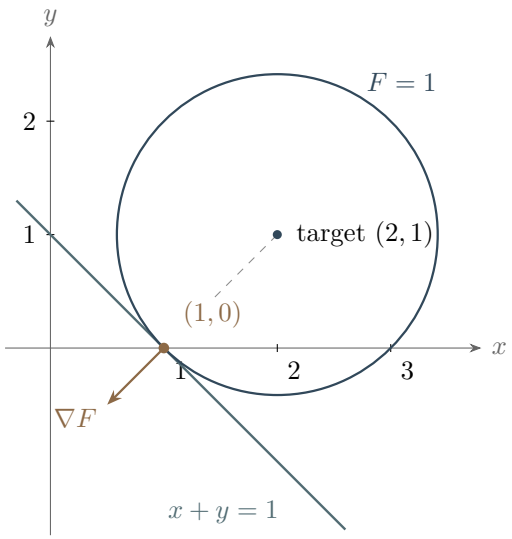


FIGURE 7.1. The contour $F = 1$ touches the feasible line at $(1, 0)$. The arrow is one half of $\nabla F(1, 0) = (-1, -1)^\mathsf{T}$. Equal axis scales preserve orthogonality.

For several C^1 equalities $g_j = c_j$, linearly independent gradients give $\nabla f = \sum_j \lambda_j \nabla g_j$. Their span is the normal space; the common kernel of their transposes is the tangent space. We retain one equality in the examples so the geometry remains explicit.

7.3 A constraint changes the fitted optimum

Example 7.2. A certified endpoint in a calibration fit

The observations $(1, 2, 4)$ in the temperature calibration are imperfect measurements. Suppose a separate reference now requires the fitted response at $t = 1$ to be exactly 3, imposing $a + b = 3$. We retain all three observations in J to quantify disagreement with the original data. This hard requirement overrides the endpoint reading 4.

With $L = J - \lambda(a + b - 3)$, stationarity gives

$$6a - 14 = \lambda, \qquad 4b - 6 = \lambda, \qquad a + b = 3.$$

Substitute $b = 3 - a$ and equate the first two expressions: $10a = 20$. Therefore

$$a = 2, \qquad b = 1, \qquad \lambda = -2, \qquad J = 1.$$

The fitted responses are $(1, 2, 3)^\mathsf{T}$, with residual $(0, 0, 1)^\mathsf{T}$. All the remaining loss is at the endpoint where the observed and certified values disagree.

For a complete optimality argument, eliminate b:

$$J(a, 3 - a) = 1 + 5(a - 2)^2. \tag{7.3}$$

The solution is the unique global constrained minimum, with loss penalty $1 - 1/6 = 5/6$ relative to the unconstrained fit. Its gradient $(-2, -2)^\mathsf{T}$ is nonzero but perpendicular to the feasible direction $(1, -1)^\mathsf{T}$. Moving along the negative gradient would break the certified requirement.

7.4 Allocation and equal marginal costs

Example 7.3. Sharing a divisible processing load

A batch of data represents 6 normalized workload units, divisible between two processors. Assign amounts $x, y \geq 0$ with $x + y = 6$. Suppose estimated total processing energy, in fixed reference units, is

$$C(x, y) = x^2 + 2y^2.$$

The coefficients specify a synthetic energy model over the stated workload range; the second processor has a larger energy cost for the same load.

At an interior allocation,

$$2x = \lambda, \qquad 4y = \lambda, \qquad x + y = 6,$$

so $x = 4$, $y = 2$, $\lambda = 8$, and $C = 24$. An equal split costs 27. The derivatives $2x$ and $4y$ are *marginal energy costs*, per additional workload unit. They agree at the optimum. If they differed, moving a small amount of work from the larger margin to the smaller would save energy while preserving total load.

The full feasible set includes endpoints. Substitution gives

$$C(x, 6 - x) = 24 + 3(x - 4)^2, \qquad 0 \leq x \leq 6.$$

Thus $(4, 2)$ is the unique global solution; $(0, 6)$ and $(6, 0)$ cost 72 and 36. The equal-marginal rule selects the split, while the restricted objective certifies it over all admissible allocations.

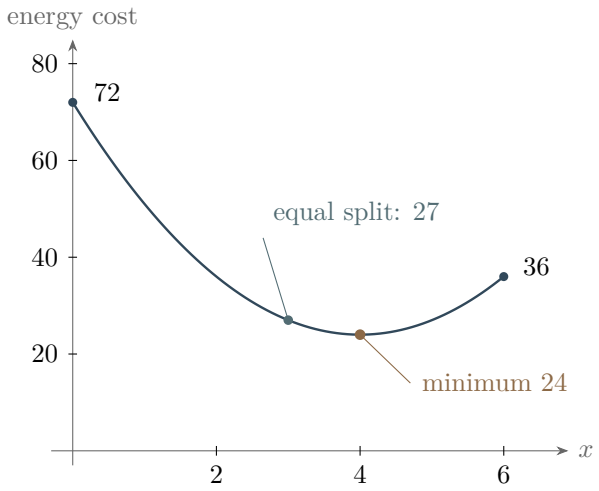


FIGURE 7.2. Energy cost on the entire feasible segment, with $y = 6 - x$. Equal workloads do not minimize this cost; equal marginal costs do. Endpoint values are included explicitly.

7.5 Curved constraints and global comparison

Example 7.4. Reinforcement and cancellation of two components

Let x, y be real normalized amplitudes of two components, with fixed total input energy $x^2 + y^2 = 1$. Combining them by addition gives squared amplitude

$$(x + y)^2 = 1 + 2xy.$$

To find the strongest reinforcement and cancellation, we therefore maximize and minimize the cross term $f(x, y) = xy$ on the unit circle.

For $g = x^2 + y^2$, the multiplier equations are

$$y = 2\lambda x, \qquad x = 2\lambda y, \qquad x^2 + y^2 = 1.$$

Neither component can be zero in a solution: the first two equations would force both to vanish. Substitution yields $4\lambda^2 = 1$. For $\lambda = 1/2$, $x = y$, giving

$$(x, y) = (1/\sqrt{2}, 1/\sqrt{2}), \quad (-1/\sqrt{2}, -1/\sqrt{2}), \qquad f = 1/2.$$

For $\lambda = -1/2$, $x = -y$, giving

$$(x, y) = (1/\sqrt{2}, -1/\sqrt{2}), \quad (-1/\sqrt{2}, 1/\sqrt{2}), \qquad f = -1/2.$$

The circle is closed and bounded, hence compact, and f is continuous, so both extrema exist. Its constraint gradient $2(x, y)^\mathsf{T}$ is nonzero everywhere; every extremum must be among these candidates. Comparing their values gives global bounds $-1/2 \le xy \le 1/2$.

Independently, $(|x| - |y|)^2 \ge 0$ gives $2|xy| \le x^2 + y^2 = 1$. The combined squared amplitude therefore ranges from 0 to 2. Opposite equal components cancel; matching equal components reinforce.

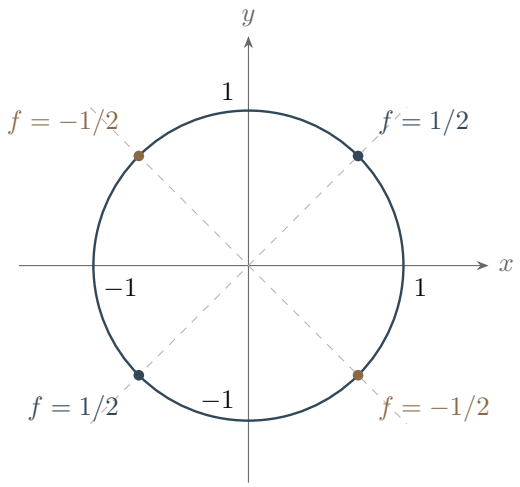


FIGURE 7.3. The four extrema of xy on $x^2 + y^2 = 1$. Each marked coordinate has magnitude $1/\sqrt{2}$; equal axis units preserve the circle and diagonal directions.

7.6 Boundaries and singular constraints

Example 7.5. A capacity limit activates a boundary

Two services request normalized processing rates $(2, 1)$, but available capacity is only $1/2$. We choose nonnegative delivered rates closest to the requests, using

$$F(x, y) = \tfrac{1}{2}((x-2)^2 + (y-1)^2), \qquad x + y = \tfrac{1}{2}, \quad x, y \geq 0.$$

This has the same distance formula as Example 7.1, but a different physical domain. Ignoring nonnegativity yields $(3/4, -1/4)$, an impossible allocation.

Eliminate y while retaining its restrictions:

$$F(x, \tfrac{1}{2} - x) = x^2 - \tfrac{3}{2}x + \tfrac{17}{8}, \qquad 0 \leq x \leq \tfrac{1}{2}.$$

The derivative $2x - 3/2$ is negative throughout the interval. The unique minimum is therefore $(1/2, 0)$, with cost $13/8$; the other endpoint costs $17/8$. Under this equal squared-deviation objective, all capacity goes to the first service. A guaranteed minimum service rate would be an additional constraint.

At the answer the inequality $y \geq 0$ is active, so the equality-only tangent argument does not apply. Substitution and endpoint comparison solve this case without a general theory of inequality multipliers.

Singularity is a different obstruction. For $f(x, y) = x$ subject to $g(x, y) = x^2 + y^2 = 0$, the only feasible point is $(0, 0)$, hence both a global minimum and a global maximum on that set. Yet $\nabla f = (1, 0)^\mathsf{T}$ and $\nabla g = \mathbf{0}$ there, so no multiplier satisfies (7.2). The equation $\nabla g^\mathsf{T} \boldsymbol{v} = 0$ admits every vector, while every actual feasible path is constant. Regularity is what rules out this discrepancy.

7.7 Further perspective: the meaning of a multiplier

Example 7.6. How an updated reference changes the best fit

Let the required endpoint response in the calibration be c, so $a + b = c$. If the reference value is updated, how much does the *best attainable* loss change after refitting? Solving the multiplier equations gives

$$a_*(c) = \frac{2c+4}{5}, \qquad b_*(c) = \frac{3c-4}{5}, \qquad \lambda(c) = \frac{12c-46}{5}.$$

Substitution gives $J(a, c-a) = V(c) + 5(a - a_*(c))^2$, where

$$V(c) = \frac{1}{6} + \frac{6}{5}\left(c - \frac{23}{6}\right)^2. \tag{7.4}$$

Thus these parameters minimize the loss globally for each c. The unconstrained endpoint prediction is 23/6; prescribing that value adds no loss.

Differentiation yields $V'(c) = \lambda(c)$. At $c = 3$, $\lambda = -2$: increasing the required response slightly lowers optimal loss initially by 2 loss units per response unit. For an update $\delta c = 0.1$,

$$V(3 + \delta c) = 1 - 2\delta c + \tfrac{6}{5}(\delta c)^2 = 0.812,$$

whereas the linear prediction is 0.8. The multiplier quantifies sensitivity to a changed requirement; it does not make the certified reference itself a free fit parameter.

Reasoning: derive the value sensitivity with the chosen sign. Suppose $f, g \in C^1$, the minimizing branch $\boldsymbol{x}_*(c)$ is differentiable, $g(\boldsymbol{x}_*(c)) = c$, and $\nabla f(\boldsymbol{x}_*) = \lambda \nabla g(\boldsymbol{x}_*)$. Assume f, g have no further explicit dependence on c. Then

$$V'(c) = \nabla f(\boldsymbol{x}_*)^{\mathsf{T}} \boldsymbol{x}_*'(c) = \lambda\, \nabla g(\boldsymbol{x}_*)^{\mathsf{T}} \boldsymbol{x}_*'(c) = \lambda,$$

because differentiation of the constraint makes the final scalar product 1. □

The quadratic example supplies the required smooth optimizer explicitly. Such a branch need not exist in a general constrained problem.

Curvature along a curved constraint

Assume $f, g \in C^2$ near a regular multiplier candidate $\boldsymbol{p}$. The relevant second-order matrix is $H_{xx}L = H_f - \lambda H_g$, restricted to $T_{\boldsymbol{p}} = \ker(\nabla g(\boldsymbol{p})^\mathsf{T})$.

Reasoning: include the curvature of the feasible path. For a C^2 feasible path through $\boldsymbol{p}$, write $\boldsymbol{v} = \boldsymbol{x}'(0)$ and $\boldsymbol{a} = \boldsymbol{x}''(0)$. Twice differentiating the constraint gives $\boldsymbol{v}^\mathsf{T} H_g \boldsymbol{v} + \nabla g^\mathsf{T} \boldsymbol{a} = 0$. Since $\nabla f = \lambda \nabla g$ at the candidate,

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}s^2} f(\boldsymbol{x}(s))\right|_{s=0} = \boldsymbol{v}^\mathsf{T} H_f \boldsymbol{v} + \nabla f^\mathsf{T} \boldsymbol{a} = \boldsymbol{v}^\mathsf{T} (H_f - \lambda H_g) \boldsymbol{v}.$$

All quantities on the right are evaluated at $\boldsymbol{p}$. The multiplier relation eliminates the path acceleration. □

Nonnegative values of this quadratic form on $T_{\boldsymbol{p}}$ are necessary for a constrained minimum. Strictly positive values for every nonzero tangent vector are sufficient for a strict local minimum under these regular C^2 assumptions; strictly negative values give a strict local maximum. Semidefinite tangent curvature with a nonzero zero-curvature direction is inconclusive.

For example, maximize the component $f(x, y) = x$ under unit energy $g = x^2 + y^2 = 1$. At $(1, 0)$, $\lambda = 1/2$, the tangent direction is vertical, and $H_f - \lambda H_g = -I$. The point is a strict constrained local maximum although $H_f = 0$. The bound $x \leq 1$ on the circle also proves global maximality. The constraint's curvature supplies the second-order decrease that the linear objective alone lacks.

7.8 Questions and exercises

Questions for review

T7.1. Define a feasible set and a constrained local extremum. Why may the ordinary gradient be nonzero at a constrained optimum?

T7.2. Derive the tangent condition $\nabla g^{\mathsf{T}} \boldsymbol{v} = 0$ from the chain rule. What extra conclusion does regularity supply?

T7.3. Explain the multiplier condition using orthogonal complements. State its assumptions and the sign convention for L used in this text.

T7.4. Describe how to turn multiplier candidates into a justified answer. What role can compactness, direct substitution, or an inequality play?

T7.5. Explain the equal-marginal-cost condition in a fixed-workload problem. Why should endpoints still be checked when workloads must be nonnegative?

T7.6. Distinguish failure caused by an active inequality boundary from failure caused by a singular equality constraint. Give one example of each.

Exercises

7.1. Find the point on $x + y = 2$ closest to $(3, 1)$. Use both a Lagrange system and a one-variable parametrization, and verify that the correction is perpendicular to the feasible line.

7.2. Minimize the recurring calibration loss $J(a, b)$ subject to $a + b = 4$. Find the parameters, multiplier, and minimum loss. Prove global optimality after eliminating b.

7.3. Minimize $C(x, y) = 2x^2 + y^2$ subject to $x + y = 9$, $x, y \geq 0$. Interpret the derivative balance, check both endpoints, and compare with an equal split.

7.4. Find all extrema of $x + 2y$ on $x^2 + y^2 = 1$ using multipliers. Verify the global bounds using Cauchy–Schwarz.

7.5. Find the point of $x + y = 1$, $x, y \geq 0$, closest to $(2, -1)$. First compute the equality-only candidate, then solve the problem on the actual segment. Explain why the endpoint answer need not satisfy the equality-only multiplier equations.

7.6. For $f(x, y) = y$ and constraint $x^2 + y^2 = 0$, identify all feasible extrema and test the multiplier equations. State precisely which regularity assumption fails.

7.7. (*) Derive $a_*(c)$, $b_*(c)$, and $V(c)$ in Example 7.6. At $c = 4$, predict the change in optimal loss for a small increase δc and compare with the exact finite-change expression.

7.8. (*) Minimize and maximize $f(x, y) = y$ on the unit circle. At $(0, 1)$ and $(0, -1)$, calculate λ, a tangent direction, and the quadratic form of $H_f - \lambda H_g$ on that direction. Compare the classification with the elementary bound $-1 \leq y \leq 1$.

From parameter geometry to function geometry

The same fit now has two descriptions: orthogonal projection in Part I, and a convex quadratic objective in Part II. The residual is orthogonal to the model space because the loss derivative vanishes in every parameter direction. Constraints replace those directions by feasible tangents.

Part III replaces sums over samples by integrals. For continuous real functions $s, \phi_1, \ldots, \phi_n$ on a finite interval $[\alpha, \beta]$, $\alpha < \beta$, let the model be $m(t) = \sum_j c_j \phi_j(t)$ and define

$$E(\boldsymbol{c}) = \int_\alpha^\beta \left(s(t) - \sum_{j=1}^{n} c_j \phi_j(t) \right)^2 \mathrm{d}t.$$

Expanding the finite sum before differentiating gives

$$E(\boldsymbol{c}) = \boldsymbol{c}^\mathsf{T} G \boldsymbol{c} - 2\boldsymbol{d}^\mathsf{T} \boldsymbol{c} + \int_\alpha^\beta s(t)^2 \,\mathrm{d}t, \qquad G_{ij} = \int_\alpha^\beta \phi_i(t)\phi_j(t)\,\mathrm{d}t, \quad d_i = \int_\alpha^\beta s(t)\phi_i(t)\,\mathrm{d}t.$$

Thus $\nabla E = 2(G\boldsymbol{c} - \boldsymbol{d})$ and $H_E = 2G$. As before,

$$\boldsymbol{h}^\mathsf{T} G \boldsymbol{h} = \int_\alpha^\beta \left(\sum_j h_j \phi_j(t) \right)^2 \mathrm{d}t \geq 0.$$

An orthonormal family makes $G = I$, so the optimal coefficients are scalar products $c_j = d_j$. This is the route to Fourier approximation: change the space of observations, retain the projection and optimization arguments.

Integration and changes of variables will also use the Jacobian DF as the local matrix of a coordinate map. For a regular change between spaces of the same dimension, $|\det DF|$ describes local volume scaling. Later approximation and transform calculations build on two tools established here: linear derivatives for local changes, and Gram quadratic forms for approximation.