Applied Optimization: Formulation and Algorithms for Engineering Systems Slides

Ross Baldick
Department of Electrical and Computer Engineering
The University of Texas at Austin
Austin, TX 78712

Part III

Unconstrained optimization

9

Case studies of unconstrained optimization

(i) Multi-variate linear regression (Section 9.1), and
(ii) State estimation in an electric power system (Section 9.2).

9.1 Multi-variate linear regression

9.1.1 Motivation

- Suppose we have a hypothesized functional relationship between **dependent variables** that vary according to some function of some **independent variables**.
- We do not have a complete specification of the function relating the variables.
- For example, if the hypothesized function is linear, the entries in coefficient matrix will typically be unknown to us.
- These unknown entries are called the **parameters** of the function.

9.1.2 Formulation

9.1.2.1 Measurement variables

- Assume that there is one dependent variable in our problem and call it ζ.
- Also assume that there are $(n-1)$ independent variables.
- Collect the independent variables together into a vector $\psi \in \mathbb{R}^{n-1}$.

9.1.2.2 Functional relationship

- We believe that there is an affine relationship between ζ and ψ.

$$\forall \psi \in \mathbb{R}^{n-1}, \zeta = \beta^{\dagger}\psi + \gamma. \tag{9.1}$$

- We want to find the unknowns in the vector $x = \begin{bmatrix} \beta \\ \gamma \end{bmatrix} \in \mathbb{R}^n$.

9.1.2.3 Trials

- We can perform a number of "trials" with varying values for the independent variables ψ.
- We use $\psi(\ell)$ and $\zeta(\ell)$, respectively, to denote the value of the independent variables ψ and the corresponding measured value of the dependent variable ζ for the ℓ-th trial.

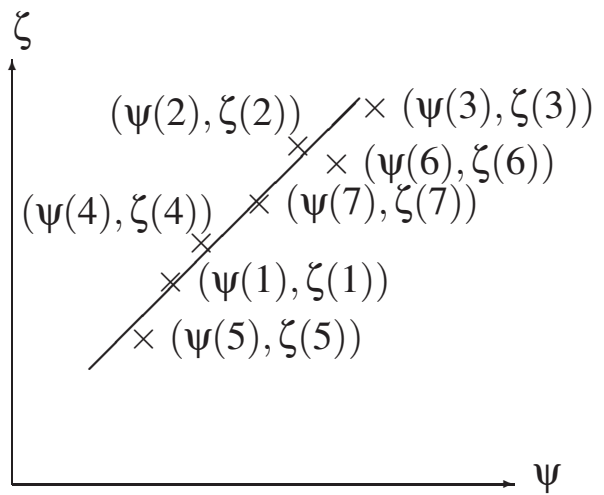

Fig. 9.1. The values of $(\psi(\ell), \zeta(\ell))$ (shown as $\times$) and affine fit.

9.1.2.4 Measurement error

$$\zeta(\ell) = \beta^\dagger \psi(\ell) + \gamma + e_\ell. \tag{9.2}$$

- The measurement **error** e_ℓ is also called the **residual**.

Calibration error

- There may be a function $c : \mathbb{R} \to \mathbb{R}$, called the **calibration function**, such that:

$$\beta^\dagger \psi(\ell) + \gamma = \zeta(\ell) - c(\zeta(\ell)).$$

Functional error

- The error e_ℓ may be due to error in the assumed functional form:

$$\zeta = \beta^\dagger \psi + \psi^\dagger \Gamma \psi, \tag{9.3}$$

- where $\Gamma \in \mathbb{R}^{(n-1)\times(n-1)}$ is a matrix of unknown parameters.

Random error

- The error e_ℓ may be **random** with expected value, say, 0.
- That is, ζ also depends on other variables besides ψ that we can neither control nor measure easily.
- It may be reasonable to model these errors as random variables that vary independently of the trials as in the following examples.

Black-box circuit

- It may be reasonable to assume that the temperature is independent of the injected currents.

Drug efficacy

- It may be reasonable to assume that immune system properties vary randomly from patient to patient and are independent of the symptoms, drugs, and treatment.

Discussion

- We should be very cautious about asserting independence between the independent variables $\psi(\ell)$ and the error e_ℓ.

9.1.2.5 Random error distribution

- We will only consider random error in this case study.

Central limit theorem

- Suppose that there are a number of factors that sum to e_ℓ in trial ℓ.
- The **central limit theorem** says that the sum of a large number of independent random variables has a distribution that is approximately **Gaussian**, with density:

$$\frac{1}{\sqrt{2\pi}\sigma_\ell}\exp\left(-\frac{(e_\ell-\mu_\ell)^2}{2(\sigma_\ell)^2}\right), \tag{9.4}$$

- where μ_ℓ is the expected value of e_ℓ, in our case 0, and σ_ℓ is its standard deviation.

Error correlation

- We will assume that e_ℓ is uncorrelated with $e_{\ell'}$ for $\ell \neq \ell'$.

Distribution of dependent variables

$$\forall \zeta(\ell) \in \mathbb{R}, \phi_\ell(\zeta(\ell); \psi(\ell), x) = \frac{1}{\sqrt{2\pi}\sigma_\ell} \exp\left(-\frac{(\zeta(\ell) - \beta^\dagger \psi(\ell) - \gamma)^2}{2(\sigma_\ell)^2} \right).$$

- We use a semi-colon to separate the arguments of the function from the parameters $\psi(\ell)$ and x.

Joint measurement distribution

- If the error distributions are **jointly Gaussian** and uncorrelated then the joint probability density function, $\phi : \mathbb{R}^m \to \mathbb{R}$, is the product of the individual probability densities:

$$\begin{aligned} \forall \zeta(1) \in \mathbb{R}, \ldots, \forall \zeta(m) \in \mathbb{R}, \phi(\zeta(1), \ldots, \zeta(m); \psi(1), \ldots, \psi(m), x) \\ = \prod_{\ell=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_\ell} \exp\left(-\frac{(\zeta(\ell) - \beta^\dagger \psi(\ell) - \gamma)^2}{2(\sigma_\ell)^2} \right). \end{aligned} \quad (9.5)$$

9.1.2.6 *Problem variables*

- *After* performing the trials, the values of $\psi(\ell)$ and $\zeta(\ell)$ are known and we will re-interpret them as constants.
- The unknowns are the parameters β and γ in the relationship (9.1).
- We have collected together these parameters into the vector x and they will be re-interpreted as the variables in our problem formulation since they are the values that are to be determined to solve our regression problem.

9.1.2.7 Maximum likelihood estimation

- Need a criterion for choosing the "best value."
- Suppose that we are given:
  - a collection of measurements $\zeta(1) \in \mathbb{R}, \ldots, \zeta(m) \in \mathbb{R}$,
  - values of the parameters $x \in \mathbb{R}^n$, and
  - a distance $\delta \in \mathbb{R}_+$.
- Suppose we take new measurements, $\tilde{\zeta}(1), \ldots, \tilde{\zeta}(m)$ using the same values of the independent variables
- Consider the probability that the new measurements $\tilde{\zeta}(1), \ldots, \tilde{\zeta}(m)$ lie in the set:

$$\mathbb{S}(x) = \{\tilde{\zeta}(1) \in \mathbb{R}, \ldots, \tilde{\zeta}(m) \in \mathbb{R} | \zeta(\ell) - \delta \leq \tilde{\zeta}(\ell) \leq \zeta(\ell) + \delta, \forall \ell = 1, \ldots, m\}.$$

- This probability is approximately equal to:

$$\phi(\zeta(1), \ldots, \zeta(m); \psi(1), \ldots, \psi(m), x)(2\delta)^m.$$

- We pick $x \in \mathbb{R}^n$ to maximize the probability that the new measurements are in the set $\mathbb{S}(x)$, which is equivalent to maximizing:

$$\phi(\zeta(1),\ldots,\zeta(m);\psi(1),\ldots,\psi(m),x)$$

- over $x \in \mathbb{R}^n$.
- We now maximize ϕ, *re-interpreted* to be the function $\phi : \mathbb{R}^n \to \mathbb{R}$ defined by:

$$\begin{aligned}
&\forall x \in \mathbb{R}^n, \phi(\zeta(1),\ldots,\zeta(m);\psi(1),\ldots,\psi(m),x) \\
&\qquad = \prod_{\ell=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_\ell} \exp\left(-\frac{(\zeta(\ell)-\beta^\dagger\psi(\ell)-\gamma)^2}{2(\sigma_\ell)^2}\right), \\
&\qquad = \prod_{\ell=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_\ell} \exp\left(-\frac{(\psi(\ell)^\dagger\beta+\gamma-\zeta(\ell))^2}{2(\sigma_\ell)^2}\right), \qquad (9.6)
\end{aligned}$$

- where $x = \begin{bmatrix} \beta \\ \gamma \end{bmatrix}$.

9.1.2.8 Problem

- The maximum likelihood estimation problem:

$$\max_{x \in \mathbb{R}^n} \phi(\zeta(1), \ldots, \zeta(m); \psi(1), \ldots, \psi(m), x). \tag{9.7}$$

9.1.3 Change of number of trials or correction of data

- We may find that after solving the maximum likelihood estimation using trials $1, \ldots, m$ we conduct further trials or find that some of the data is in error and needs to be corrected.
- We would like to be able obtain an updated estimation without starting from scratch.

9.1.4 Problem characteristics

9.1.4.1 Parameters re-interpreted as variables

- We have re-interpreted the *parameters* β and γ of the probability density in (9.5) to be the *variables* in our optimization problem.
- We interpret $\psi(\ell)$ and $\zeta(\ell)$ to be *known* values once the trials have been completed.

9.1.4.2 Objective

- The objective $\phi(\zeta(1),\ldots,\zeta(m);\psi(1),\ldots,\psi(m),x)$ is the product of terms.
- Each term in the product depends on x.

9.1.4.3 Number of parameters and trials

- If $m \leq n$ then there is no redundancy and we will not be able to reduce the effects of measurement errors.

9.1.4.4 Generalizations

- In some cases, we may have a non-linear relationship between the dependent and independent variables, as in (9.3).

9.2 Power system state estimation

- We formulate a **non-linear regression** problem.

9.2.1 Motivation

9.2.1.1 Non-linear regression

- Suppose that we hypothesize a non-linear relationship, such as $\zeta = \gamma(\psi)^{\beta}$, between scalars ψ and ζ with unknown parameters β and γ.
- A standard approach for this particular non-linear relationship is to take logarithms of both sides to form the equation:

$$\begin{aligned} \ln(\zeta) &= \beta \ln(\psi) + \ln(\gamma), \\ \Psi &= \ln(\psi), \\ Z &= \ln(\zeta). \end{aligned}$$

- We have implicitly defined an onto function $\tau : \mathbb{R}^2_{++} \to \mathbb{R}^2$ and a transformed functional relationship specified by:

$$\forall \begin{bmatrix} \psi \\ \zeta \end{bmatrix} \in \mathbb{R}^2_{++}, \tau\left(\begin{bmatrix} \psi \\ \zeta \end{bmatrix}\right) = \begin{bmatrix} \ln(\psi) \\ \ln(\zeta) \end{bmatrix},$$
$$Z = \beta\Psi + \Gamma.$$

- It is not always possible to find such a transformation.
- For example, consider a functional relationship between scalars ψ and ζ of the form:

$$\zeta = \gamma(\psi)^\beta + \delta\psi.$$

- We cannot transform this equation in a way such that all the unknown parameters β, γ, and δ (or their transformed versions) appear linearly.
- Such a problem is called a **non-linear regression** problem.

9.2.1.2 *Power system measurements*

- We may want to observe the *actual* state of the system to check if the system is operating within limits.
- The state estimation problem involves finding the voltage angles and magnitudes in the system that best match the measured values.

9.2.2 Formulation

9.2.2.1 Measurements

- Real and reactive power injection at a bus;
- Real and reactive power flow along a line; and
- Voltage magnitude.

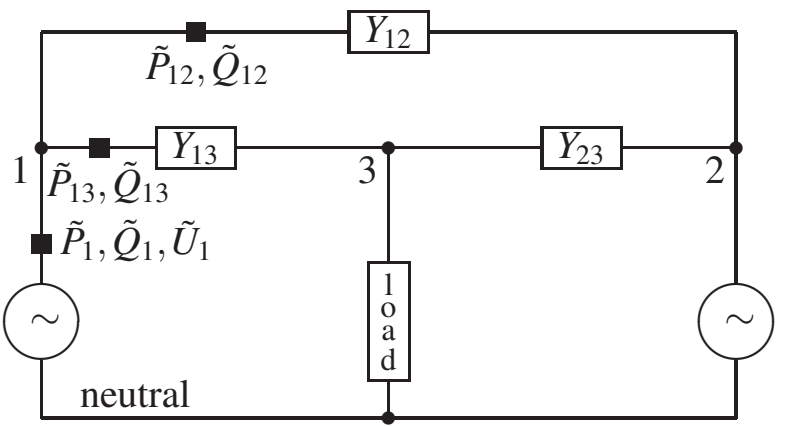

Fig. 9.2. Three-bus power system state estimation problem.

Real and reactive power injection

- Let $\mathbb{B}$ be the set of buses where there are measurements of the real and reactive power injections into the system.
- In Figure 9.2, $\mathbb{B} = \{1\}$.

Real and reactive line flow

- Let $\mathbb{F}$ be the set of lines where we have line flow measurements.
- In Figure 9.2, $\mathbb{F} = \{(1,2),(1,3)\}$.

Voltage magnitude

- Finally, let $\mathbb{U}$ be the set of buses where there are voltage magnitude measurements.
- In Figure 9.2, $\mathbb{U} = \{1\}$.

9.2.2.2 Variables

- Change the definition of x in Section 6.2 to include:
  - the voltage angles at all buses except the reference bus, and
  - the voltage magnitudes at all buses in the system, including the reference bus.
- Now $x \in \mathbb{R}^n$, where n is equal to one less than twice the number of buses, so that the vector x has been re-defined compared to Section 6.2.

9.2.2.3 Measurement functions

- Recall the definitions of the functions $p_\ell, q_\ell : \mathbb{R}^n \to \mathbb{R}$ in (6.12) and (6.13) that were used in the power flow case study:

$$\forall x \in \mathbb{R}^n, p_\ell(x) = \sum_{k \in \mathbb{J}(\ell) \cup \{\ell\}} u_\ell u_k [G_{\ell k} \cos(\theta_\ell - \theta_k) + B_{\ell k} \sin(\theta_\ell - \theta_k)] - P_\ell,$$

$$\forall x \in \mathbb{R}^n, q_\ell(x) = \sum_{k \in \mathbb{J}(\ell) \cup \{\ell\}} u_\ell u_k [G_{\ell k} \sin(\theta_\ell - \theta_k) - B_{\ell k} \cos(\theta_\ell - \theta_k)] - Q_\ell.$$

- Let us define new functions by omitting the values of the real and reactive injections, P_ℓ and Q_ℓ.
- That is, define $\tilde{p}_\ell : \mathbb{R}^n \to \mathbb{R}$ and $\tilde{q}_\ell : \mathbb{R}^n \to \mathbb{R}$ to be:

$$\forall x \in \mathbb{R}^n, \tilde{p}_\ell(x) = \sum_{k \in \mathbb{J}(\ell) \cup \{\ell\}} u_\ell u_k [G_{\ell k} \cos(\theta_\ell - \theta_k) + B_{\ell k} \sin(\theta_\ell - \theta_k)],$$

$$\forall x \in \mathbb{R}^n, \tilde{q}_\ell(x) = \sum_{k \in \mathbb{J}(\ell) \cup \{\ell\}} u_\ell u_k [G_{\ell k} \sin(\theta_\ell - \theta_k) - B_{\ell k} \cos(\theta_\ell - \theta_k)].$$

9.2.2.4 Measurement functions

- We denote the measurement functions by:

$$\begin{aligned} \tilde{p}_\ell, \tilde{q}_\ell, &\quad \text{for the real and reactive power injection measurements, } \ell \in \mathbb{B}, \\ \tilde{p}_{\ell k}, \tilde{q}_{\ell k}, &\quad \text{for the real and reactive line flow measurements, } (\ell, k) \in \mathbb{F}, \\ \tilde{u}_\ell, &\quad \text{for the voltage magnitude measurements, } \ell \in \mathbb{U}. \end{aligned}$$

- We collect the measurement functions into a vector function $\tilde{g}$ and collect the measurements together into a corresponding vector $\tilde{G}$:

$$\forall x \in \mathbb{R}^n, \tilde{g}(x) = \begin{pmatrix} \begin{bmatrix} \tilde{p}_\ell(x) \\ \tilde{q}_\ell(x) \end{bmatrix}_{\ell \in \mathbb{B}} \\ \begin{bmatrix} \tilde{p}_{\ell k}(x) \\ \tilde{q}_{\ell k}(x) \end{bmatrix}_{(\ell,k) \in \mathbb{F}} \\ \left[\tilde{u}_\ell(x) \right]_{\ell \in \mathbb{U}} \end{pmatrix}, \tilde{G} = \begin{pmatrix} \begin{bmatrix} \tilde{P}_\ell \\ \tilde{Q}_\ell \end{bmatrix}_{\ell \in \mathbb{B}} \\ \begin{bmatrix} \tilde{P}_{\ell k} \\ \tilde{Q}_{\ell k} \end{bmatrix}_{(\ell,k) \in \mathbb{F}} \\ \left[\tilde{U}_\ell \right]_{\ell \in \mathbb{U}} \end{pmatrix}.$$

- Let us define a new index set $\mathbb{M}$ that specifies all the measurements.
- We re-index the entries of $\tilde{g}$ and $\tilde{G}$ using the set $\mathbb{M}$, so that $\tilde{g} = (\tilde{g}_k)_{k \in \mathbb{M}}$ and $\tilde{G} \in \mathbb{R}^{\mathbb{M}}$.

9.2.2.5 Error distribution

- Assuming independent Gaussian measurement errors then we can write the probability density, $\phi : \mathbb{R}^{\mathbb{M}} \rightarrow \mathbb{R}$, of the measurement vector $\tilde{G}$ as the product of probability densities:

$$\forall \tilde{G} \in \mathbb{R}^{\mathbb{M}}, \phi(\tilde{G};x) = \prod_{\ell \in \mathbb{B}} \phi_{\tilde{p}_\ell}(\tilde{P}_\ell;x) \prod_{\ell \in \mathbb{B}} \phi_{\tilde{q}_\ell}(\tilde{Q}_\ell;x) \prod_{(\ell,k) \in \mathbb{F}} \phi_{\tilde{p}_{\ell k}}(\tilde{P}_{\ell k};x) \times \prod_{(\ell,k) \in \mathbb{F}} \phi_{\tilde{q}_{\ell k}}(\tilde{Q}_{\ell k};x) \prod_{\ell \in \mathbb{U}} \phi_{\tilde{u}_\ell}(\tilde{U}_\ell;x),$$

- where each function $\phi_{\tilde{p}_\ell}(\tilde{P}_\ell;x), \phi_{\tilde{q}_\ell}(\tilde{Q}_\ell;x), \phi_{\tilde{p}_{\ell k}}(\tilde{P}_{\ell k};x), \phi_{\tilde{q}_{\ell k}}(\tilde{Q}_{\ell k};x)$, and $\phi_{\tilde{u}_\ell}(\tilde{U}_\ell;x)$ represents the probability density function of the corresponding error distribution and is parameterized by x.

- For example,

$$\forall \tilde{P}_\ell \in \mathbb{R}, \phi_{\tilde{p}_\ell}(\tilde{P}_\ell; x) = \frac{1}{\sqrt{2\pi}\sigma_{\tilde{p}_\ell}} \exp\left(-\frac{(\tilde{p}_\ell(x) - \tilde{P}_\ell)^2}{2(\sigma_{\tilde{p}_\ell})^2}\right),$$

- where $\sigma_{\tilde{p}_\ell}$ is the standard deviation of the measurement error of real power at bus ℓ and where we have assumed that the expected error is zero.
- After the measurements are made, we can re-interpret ϕ to be a function $\phi : \mathbb{R}^n \to \mathbb{R}$. That is, we re-interpret ϕ as being defined by:

$$\begin{aligned}\forall x \in \mathbb{R}^n, \phi(\tilde{G}; x) \;=\; & \prod_{\ell \in \mathbb{B}} \phi_{\tilde{p}_\ell}(\tilde{P}_\ell; x) \prod_{\ell \in \mathbb{B}} \phi_{\tilde{q}_\ell}(\tilde{Q}_\ell; x) \prod_{(\ell,k) \in \mathbb{F}} \phi_{\tilde{p}_{\ell k}}(\tilde{P}_{\ell k}; x) \\ & \times \prod_{(\ell,k) \in \mathbb{F}} \phi_{\tilde{q}_{\ell k}}(\tilde{Q}_{\ell k}; x) \prod_{\ell \in \mathbb{U}} \phi_{\tilde{u}_\ell}(\tilde{U}_\ell; x).\end{aligned}$$

- Our maximum likelihood estimation problem is then:

$$\max_{x \in \mathbb{R}^n} \phi(\tilde{G}; x). \tag{9.8}$$

9.2.3 Change in measurement data

- We will consider how a change in measurement data affects the result.

9.2.4 Problem characteristics

9.2.4.1 Objective

- The objective of this problem is very similar to that of multi-variate linear regression Problem (9.7), except that each term in the product has one of the non-linear functions $\tilde{p}_\ell, \tilde{q}_\ell, \tilde{p}_{\ell k}, \tilde{q}_{\ell k}$, or $\tilde{u}_\ell$ in the exponent instead of the linear measurement equation $\psi(\ell)^\dagger \beta + \gamma$.

9.2.4.2 *Solvability*

- The measurements shown in the system illustrated in Figure 9.2 have just enough information to determine all the values of the entries in x.
- It is important to have redundancy of measurements in the system and to "spread out" the measurements across the system as illustrated in Figure 9.3.

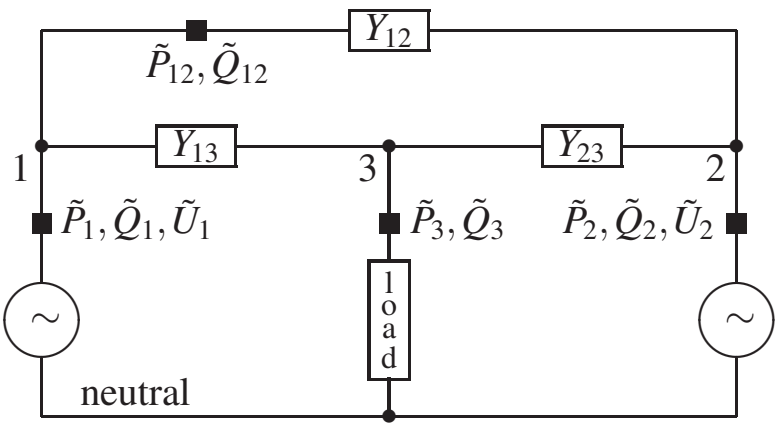

Fig. 9.3. Three-bus power system state estimation problem with spread out measurements.

10

Algorithms for unconstrained minimization

- In this chapter we will develop algorithms for unconstrained optimization problems of the form:

$$\min_{x \in \mathbb{R}^n} f(x),$$

- where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$.

Key issues

- **Descent directions** to reduce the value of the objective,
- optimality conditions based on **descent directions**,
- optimality conditions for **convex objectives**,
- the development of **iterative algorithms**, and
- **sensitivity analysis**.

Definition 10.1 Let $\hat{x} \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$. Then the vector $\Delta x \in \mathbb{R}^n$ is called a **descent direction** for f at $\hat{x}$ if:

$$\exists \overline{\alpha} \in \mathbb{R}_{++} \text{ such that } (0 < \alpha \leq \overline{\alpha}) \Rightarrow (f(\hat{x} + \alpha \Delta x) < f(\hat{x})).$$

□

10.1.1.2 Example

$$\forall x \in \mathbb{R}^2, f(x) = (x_1 - 1)^2 + (x_2 - 3)^2. \tag{10.1}$$

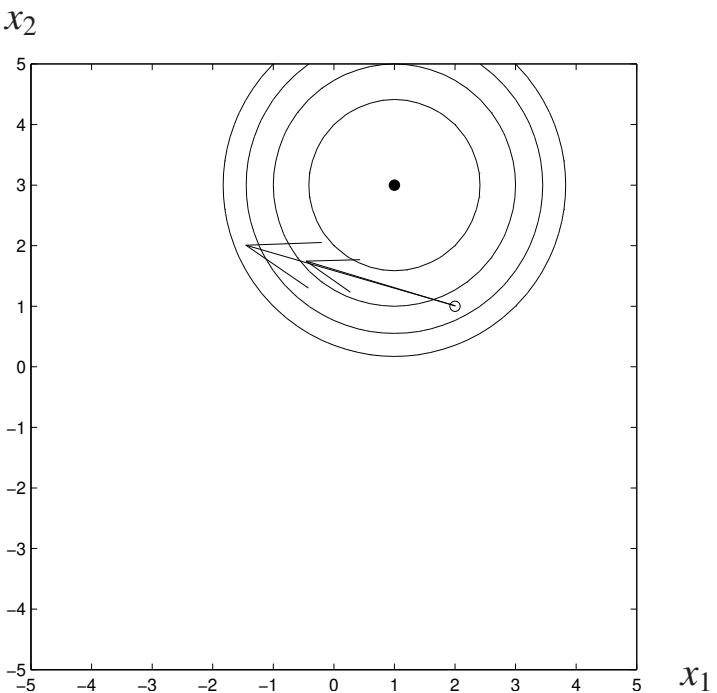

Fig. 10.1. Descent direction (shown as the longer arrow) for a function at a point $\hat{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, shown as a $\circ$. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a $\bullet$.

10.1.1.3 Steepest descent step direction

- $\Delta x = -\nabla f(x)$ is called the direction of **steepest descent**.

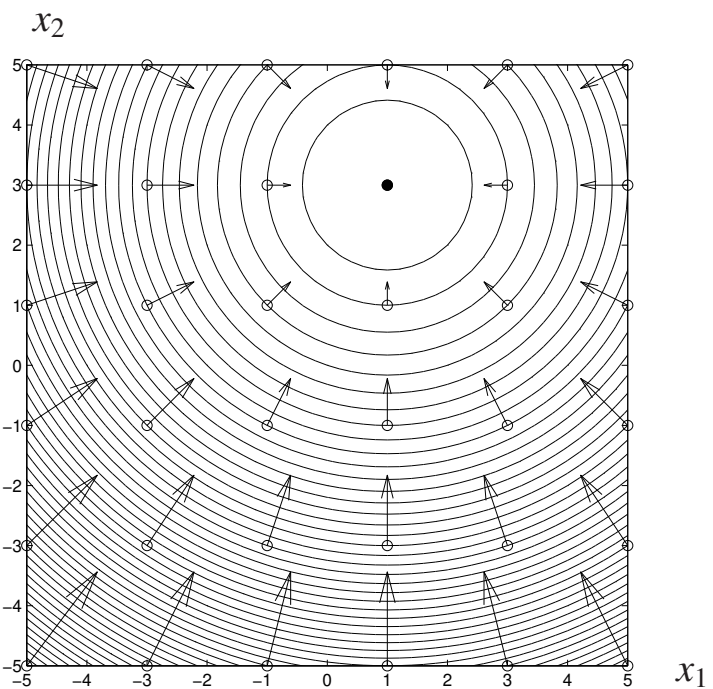

Fig. 10.2. Steepest descent directions for a function at various points. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●.

Lemma 10.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be partially differentiable with continuous partial derivatives and let $\hat{x} \in \mathbb{R}^n$, $\Delta x \in \mathbb{R}^n$. Suppose that $\nabla f(\hat{x})^\dagger \Delta x < 0$. Then Δx is a descent direction for f at $\hat{x}$.*

Proof Let $\phi : \mathbb{R} \to \mathbb{R}$ be defined by:

$$\forall t \in \mathbb{R}, \phi(t) = f(\hat{x} + t\Delta x).$$

By the chain rule, $\dfrac{d\phi}{dt}(t) = \dfrac{\partial f}{\partial x}(\hat{x} + t\Delta x)\Delta x$. Evaluating this at $t = 0$ yields:

$$\begin{aligned} \frac{d\phi}{dt}(0) &= \frac{\partial f}{\partial x}(\hat{x})\Delta x, \\ &= \nabla f(\hat{x})^\dagger \Delta x, \\ &= -2\varepsilon, \end{aligned}$$

say, where $\varepsilon > 0$ by assumption.

Proof, continued But, by definition, since f is partially differentiable with continuous partial derivatives,

$$\frac{d\phi}{dt}(0) = \lim_{\alpha \to 0} \frac{f(\hat{x} + \alpha \Delta x) - f(\hat{x})}{\alpha}.$$

Let $\overline{\alpha} \in \mathbb{R}_{++}$ be small enough such that

$$(0 < |\alpha| \leq \overline{\alpha}) \Rightarrow \left(\left| \frac{f(\hat{x} + \alpha \Delta x) - f(\hat{x})}{\alpha} - \frac{d\phi}{dt}(0) \right| \leq \varepsilon \right).$$

But this means that:

$$(0 < |\alpha| \leq \overline{\alpha}) \Rightarrow \left(\left| \frac{f(\hat{x} + \alpha \Delta x) - f(\hat{x})}{\alpha} - (-2\varepsilon) \right| \leq \varepsilon \right),$$

which implies that:

$$(0 < |\alpha| \leq \overline{\alpha}) \Rightarrow \left(\frac{f(\hat{x} + \alpha \Delta x) - f(\hat{x})}{\alpha} \leq -\varepsilon \right).$$

Proof, continued So:

$$\begin{aligned}(0 < \alpha \leq \overline{\alpha}) &\Rightarrow (f(\hat{x}+\alpha\Delta x) - f(\hat{x}) \leq -\alpha\varepsilon < 0),\\ &\Rightarrow (f(\hat{x}+\alpha\Delta x) < f(\hat{x})),\end{aligned}$$

and Δx is a descent direction for f at $\hat{x}$. □

- $\nabla f(\hat{x})^{\dagger}\Delta x$ is called the **directional derivative of f at $\hat{x}$ in the direction Δx.**
- Analytically, the condition in Lemma 10.1 that $\nabla f(\hat{x})^{\dagger}\Delta x < 0$ requires that the directional derivative in the direction Δx be negative.
- Geometrically, this condition requires that the angle between Δx and $-\nabla f(\hat{x})$ be less than 90° for Δx to be a descent direction as illustrated in Figure 10.3.

Descent directions

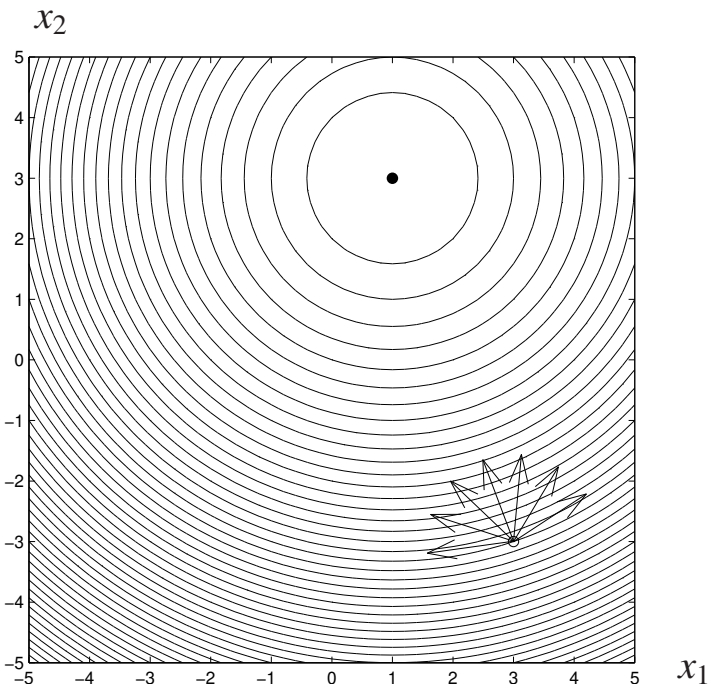

Fig. 10.3. Various descent directions for a function a particular point $\hat{x} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$. The contours decrease towards the point $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a •.

Corollary 10.2 *Let $\hat{x} \in \mathbb{R}^n$, let $M \in \mathbb{R}^{n \times n}$ be positive definite, and let $f : \mathbb{R}^n \to \mathbb{R}$ be partially differentiable with continuous partial derivatives and such that $\nabla f(\hat{x}) \neq \mathbf{0}$. Then $\Delta x = -M \nabla f(\hat{x})$ is a descent direction for f at $\hat{x}$.*

Proof Note that $\nabla f(\hat{x})^\dagger \Delta x = -\nabla f(\hat{x})^\dagger M \nabla f(\hat{x}) < 0$, since M is positive definite and $\nabla f(\hat{x}) \neq \mathbf{0}$. Apply Lemma 10.1. □

- The "middle" arrow in Figure 10.3 shows the steepest descent step direction at $\hat{x}$, corresponding to the choice $M = \mathbf{I}$.
- The other directions correspond to other choices of positive definite M and also yield descent directions in that f is also reducing in these directions away from $\hat{x}$.

10.1.2 First-order conditions

10.1.2.1 Necessary conditions

Theorem 10.3 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be partially differentiable with continuous partial derivatives. If $x^\star$ is a local minimizer of f then $\nabla f(x^\star) = \mathbf{0}$.*

Proof We prove the contra-positive. That is, we prove that if $\nabla f(x^\star) \neq \mathbf{0}$ then $x^\star$ is not a local minimizer. Let $M \in \mathbb{R}^{n\times n}$ be positive definite. By Corollary 10.2, $\Delta x = -M\nabla f(x^\star)$ is a descent direction for f at $x^\star$ and so $x^\star$ is not a local minimizer of f. □

- The statement and proof of Theorem 10.3, respectively, suggest two approaches to finding a minimizer of f:
  - (i) solve $\nabla f(x) = \mathbf{0}$, or
  - (ii) from the current point x, move in the direction $\Delta x = -M\nabla f(x)$, where M is positive definite.

10.1.2.2 Example of insufficiency

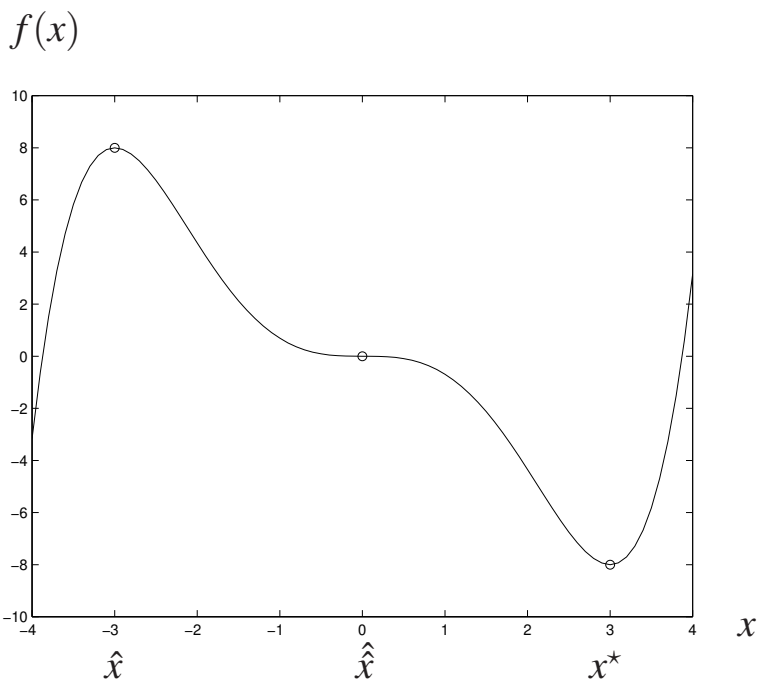

Fig. 10.4. Graph of f and points (illustrated by the $\circ$) satisfying $\nabla f(x) = \mathbf{0}$ but which may or may not correspond to a minimum.

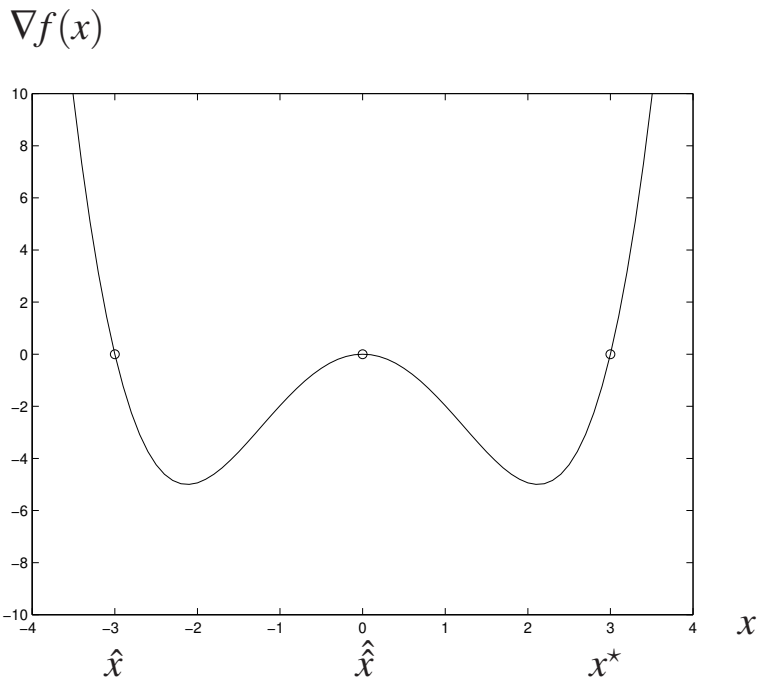

Fig. 10.5. First derivative ∇f of the function f shown in Figure 10.4.

Example of insufficiency, continued

- $\nabla f(x) = \mathbf{0}$ is not sufficient to guarantee a minimum.
- We call points that satisfy $\nabla f(x) = \mathbf{0}$ **critical points**.
- Not all critical points are minimizers.
- For the function shown in Figure 10.4:

  (i) $\hat{x} = -3, f(\hat{x}) = 8$, a local maximizer and maximum of f, respectively,
  (ii) $\hat{\hat{x}} = 0, f(\hat{\hat{x}}) = 0$, a **horizontal inflection point** of f, and
  (iii) $x^\star = 3, f(x^\star) = -8$, a local minimizer and minimum of f, respectively.

10.1.3 Second-order conditions

10.1.3.1 Necessary conditions

Analysis

Theorem 10.4 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice partially differentiable with continuous second partial derivatives and suppose that $x^\star$ is a local minimizer of f. Then:*

$$\nabla f(x^\star) = \mathbf{0}, \tag{10.2}$$

$$\nabla^2 f(x^\star) \text{ is positive semi-definite}. \tag{10.3}$$

□

Example

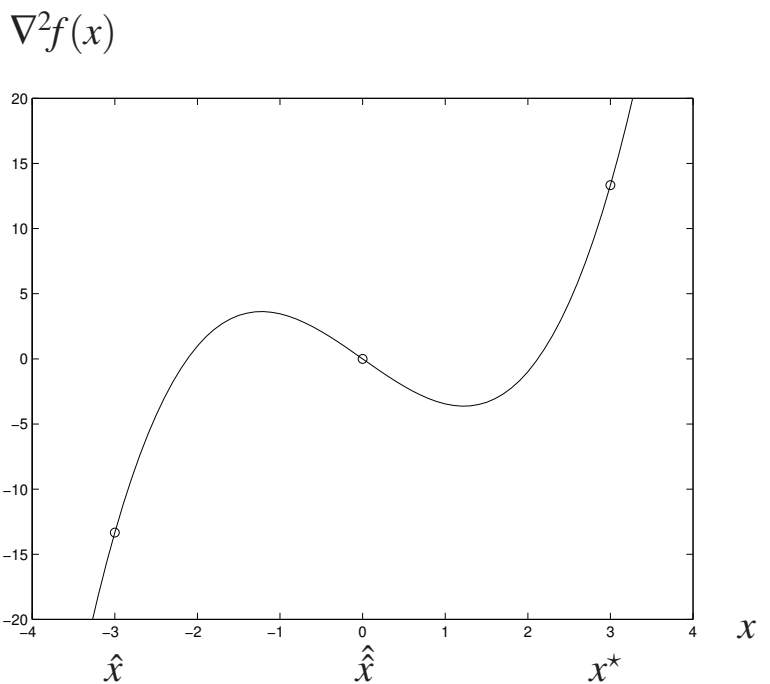

Fig. 10.6. Second derivative $\nabla^2 f$ of the function f shown in Figure 10.4.

Example, continued

- Again consider the function f shown in Figure 10.4.
- Its first and second derivatives are shown in Figures 10.5 and 10.6, respectively.
- Since $f : \mathbb{R} \to \mathbb{R}$ in this case, the Hessian $\nabla^2 f : \mathbb{R} \to \mathbb{R}$ is positive semi-definite if and only if it is non-negative.
- The critical points of f are at:

$\hat{x} = -3$. At this point, the Hessian of f, shown in Figure 10.6, is negative and hence not positive semi-definite. Therefore, by Theorem 10.4, $\hat{x} = -3$ cannot be a local minimizer of f.

$\hat{\hat{x}} = 0$. At this point, the Hessian of f is zero and hence positive semi-definite. The second-order necessary conditions are satisfied but by inspection of Figure 10.4, $\hat{\hat{x}} = 0$ is clearly not a minimizer.

$x^\star = 3$. This point is a local minimizer of f. Figure 10.6 and Theorem 10.4 both concur that the Hessian is positive semi-definite.

Analysis

Theorem 10.5 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice partially differentiable with continuous second partial derivatives and suppose that:*

$$\begin{aligned} \nabla f(x^\star) &= \mathbf{0}, \\ \nabla^2 f(x^\star) &\text{ is positive definite.} \end{aligned}$$

Then $x^\star$ is a strict local minimizer of f.

Proof By hypothesis, $\nabla^2 f(x^\star)$ is positive definite and $\nabla^2 f$ is continuous. Therefore:

$$\exists \varepsilon \in \mathbb{R}_{++} \text{ such that } (\|x^\star - x\| \leq \varepsilon) \Rightarrow (\nabla^2 f(x) \text{ is positive definite}). \tag{10.4}$$

Let Δx be any step direction such that $0 < \|\Delta x\| \leq \varepsilon$ and define $\phi : \mathbb{R} \to \mathbb{R}$ by:

$$\forall t \in \mathbb{R}, \phi(t) = f(x^\star + t\Delta x).$$

Proof, continued Then:

$$\begin{aligned}
\frac{d\phi}{dt}(t) &= \frac{\partial f}{\partial x}(x^\star + t\Delta x)\Delta x, \\
\frac{d\phi}{dt}(0) &= \frac{\partial f}{\partial x}(x^\star)\Delta x, \\
&= \nabla f(x^\star)^\dagger \Delta x, \\
&= \mathbf{0}, \text{ by hypothesis,} \qquad (10.5) \\
\frac{d^2\phi}{dt^2}(t) &= \Delta x^\dagger \frac{\partial^2 f}{\partial x^2}(x^\star + t\Delta x)\Delta x, \\
&> 0, \forall 0 < t \leq 1, \qquad (10.6)
\end{aligned}$$

where the last inequality follows from (10.4) since $\Delta x \neq 0$ and since:

$$(0 < t \leq 1) \Rightarrow (\|x^\star - (x^\star + t\Delta x)\| = t\,\|\Delta x\| \leq \|\Delta x\| \leq \varepsilon).$$

Proof, continued We have that $\phi(0) = f(x^\star)$ and:

$$
\begin{aligned}
\forall \Delta x \in \mathbb{R}^n, (0 < \|\Delta x\| \leq \varepsilon) \Rightarrow \\
f(x^\star + \Delta x) &= \phi(1), \\
&= \phi(0) + \int_{t=0}^{1} \frac{d\phi}{dt}(t)\, dt, \\
&= \phi(0) + \int_{t=0}^{1} \left[\frac{d\phi}{dt}(0) + \int_{t'=0}^{t} \frac{d^2\phi}{dt^2}(t')\, dt' \right] dt, \\
&= \phi(0) + \frac{d\phi}{dt}(0) + \int_{t=0}^{1} \int_{t'=0}^{t} \frac{d^2\phi}{dt^2}(t')\, dt'\, dt, \\
&= \phi(0) + \int_{t=0}^{1} \int_{t'=0}^{t} \frac{d^2\phi}{dt^2}(t')\, dt'\, dt, \text{ by (10.5)}, \\
&> f(x^\star), \text{ since the integrand is strictly positive by (10.6)}.
\end{aligned}
$$

That is, $x^\star$ is a strict local minimizer. □

- Positive *semi*-definiteness of the second derivative matrix at a critical point $\hat{\hat{x}}$ is *not* sufficient to guarantee that $\hat{\hat{x}}$ is a minimizer.

Example

- Continuing with the example from Section 10.1.1.2, note that:

$$\begin{aligned} \forall x \in \mathbb{R}^2, f(x) &= (x_1 - 1)^2 + (x_2 - 3)^2, \\ \forall x \in \mathbb{R}^2, \nabla^2 f(x) &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \end{aligned}$$

- which is positive definite.
- Therefore, by Theorem 10.5, the point $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is a strict local minimizer of f.

Example of insufficiency

$$\forall x \in \mathbb{R}, f(x) = -(x)^4.$$

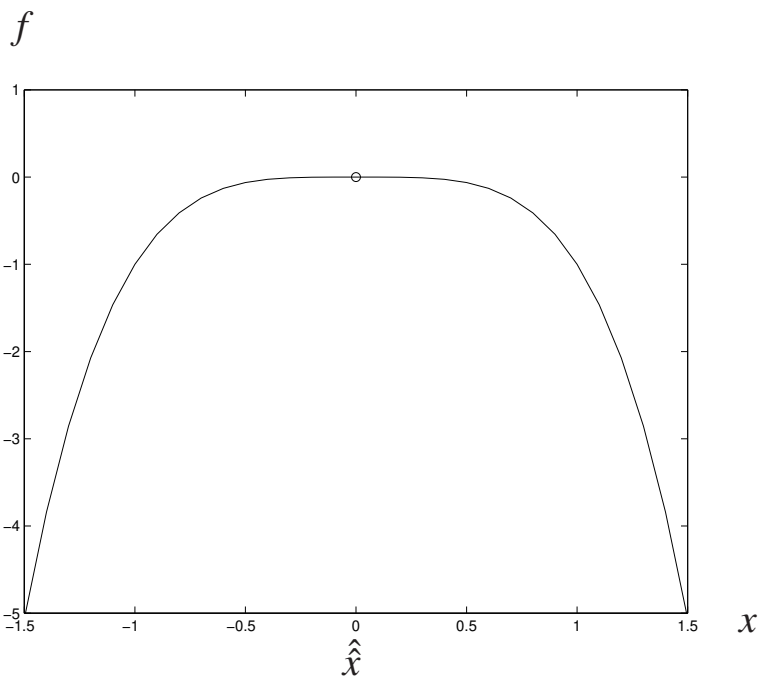

Fig. 10.7. A critical point $\hat{\hat{x}} = 0$, illustrated by the $\circ$, where the second derivative matrix is positive semi-definite at $\hat{\hat{x}}$ yet the point is not a minimizer.

Example of insufficiency, continued

- Consider the point $\hat{\hat{x}} = 0$ as illustrated in Figure 10.7.
- In this case:

$$\begin{aligned} \nabla f(\hat{\hat{x}}) &= [-4(\hat{\hat{x}})^3], \\ &= [0], \\ \nabla^2 f(\hat{\hat{x}}) &= [-12(\hat{\hat{x}})^2], \\ &= [0], \end{aligned}$$

- so that:

$$\forall \Delta x \in \mathbb{R}, 0 = \Delta x \nabla^2 f(\hat{\hat{x}}) \Delta x \geq 0,$$

- and so $\nabla^2 f(\hat{\hat{x}})$ is positive semi-definite.
- However, $\hat{\hat{x}} = [0]$ is clearly not a minimizer of f.

10.1.4 Convex objectives

10.1.4.1 First-order sufficient conditions

Analysis

- If f is twice partially differentiable with continuous partial derivatives and the second derivative matrix of f is positive semi-definite *everywhere* then the objective is convex by Theorem 2.7.

Corollary 10.6 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and partially differentiable with continuous partial derivatives on $\mathbb{R}^n$ and let $x^\star \in \mathbb{R}^n$. If $\nabla f(x^\star) = \mathbf{0}$ then $f(x^\star)$ is the global minimum and $x^\star$ is a global minimizer of f.*

Proof Recall Theorem 2.6. The hypothesis of Theorem 2.6 is satisfied for $\mathbb{S} = \mathbb{R}^n$. Consequently, (2.31) holds, which we repeat:

$$\forall x, x' \in \mathbb{S}, f(x) \geq f(x') + \nabla f(x')^\dagger (x - x').$$

Letting $x' = x^\star$ and $\mathbb{S} = \mathbb{R}^n$ in (2.31) and noting that $\nabla f(x^\star) = \mathbf{0}$, we obtain:

$$\forall x \in \mathbb{R}^n, f(x) \geq f(x^\star).$$

That is $x^\star$ is a global minimizer of f. □

Example

- Continuing with the example from Sections 10.1.1.2 and 10.1.3.2, note that $\nabla^2 f$ is positive definite so that f is convex.
- Therefore, by Corollary 10.6, the point $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is the global minimizer of f.

10.1.4.2 Uniqueness of minimizer

Theorem 10.7 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice partially differentiable with continuous second partial derivatives on $\mathbb{R}^n$. If $\nabla^2 f$ is positive definite throughout $\mathbb{R}^n$ and $\min_{x \in \mathbb{R}^n} f(x)$ possesses a minimum then the associated minimizer is unique.*

Proof Applying Theorems 2.3 and 2.2 to ∇f we find that there is at most one point that satisfies the necessary conditions for minimizing f. Alternatively, Theorem 2.7 and Item (iii) of the conclusion of Theorem 2.4 imply the same result. □

10.2 Approaches to finding minimizers

10.2.1 Steepest descent

$$x^{(\nu+1)} = x^{(\nu)} - \alpha^{(\nu)} \nabla f(x^{(\nu)}). \tag{10.7}$$

10.2.1.1 Advantages

- Unless $\nabla f(x^{(\nu)}) = \mathbf{0}$, it is always possible to find a step-size $\alpha^{(\nu)}$ such that the objective will be reduced from $f(x^{(\nu)})$ by updating the iterate to $x^{(\nu)} - \alpha^{(\nu)} \nabla f(x^{(\nu)})$.

10.2.1.2 Example

- Consider the quadratic function illustrated in Figure 10.2.

$$
\begin{aligned}
\nabla f(x) &= \begin{bmatrix} 2(x_1-1) \\ 2(x_2-3) \end{bmatrix}, \\
x^{(0)} &= \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \\
\nabla f(x^{(0)}) &= \begin{bmatrix} 2(3-1) \\ 2(-5-3) \end{bmatrix}, \\
&= \begin{bmatrix} 4 \\ -16 \end{bmatrix}, \\
x^{(1)} &= x^{(0)} + \alpha^{(0)} \Delta x^{(0)}, \\
&= \begin{bmatrix} 3 \\ -5 \end{bmatrix} + \alpha^{(0)} \begin{bmatrix} -4 \\ 16 \end{bmatrix}.
\end{aligned}
$$

- If we set $\alpha^{(0)} = 0.5$ then $x^{(1)} = x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ so that we would have reached the minimizer in one iteration.

10.2.1.3 Disadvantages

- Progress towards the solution may be very slow if the contour sets of the function are very "eccentric."

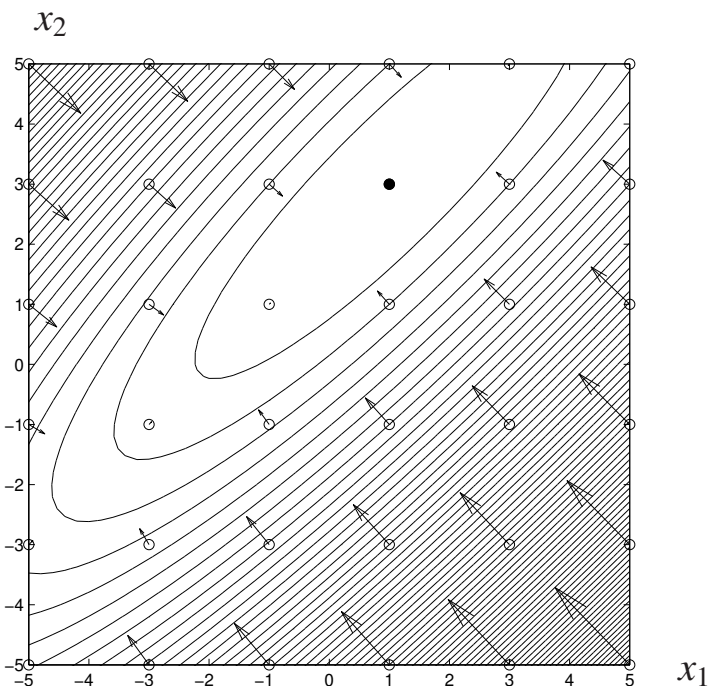

Fig. 10.8. Scaled versions of the steepest descent step directions for an objective, defined in (10.8), with contour sets that are highly eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●.

10.2.1.4 Example

- Figure 10.8 shows scaled versions of the steepest descent step directions for a quadratic function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by:

$$\begin{aligned} \forall x \in \mathbb{R}^2, f(x) &= (x_1 - 1)^2 + (x_2 - 3)^2 - 1.8(x_1 - 1)(x_2 - 3), \quad (10.8) \\ &= \frac{1}{2} x^\dagger Q x + c^\dagger x + \text{ constant}, \\ Q &= \nabla^2 f(x), \\ &= \begin{bmatrix} 2 & -1.8 \\ -1.8 & 2 \end{bmatrix}, \\ c &= \begin{bmatrix} 3.4 \\ -4.2 \end{bmatrix}. \end{aligned}$$

- This function has the same minimizer, $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, as the function in Figure 10.2, but has eccentric contour sets.
- This function is more typical of functions encountered in practice.

Example, continued

- For a step-size of $\alpha^{(\nu)}$, the next iterate has objective value:

$$f(x^{(\nu+1)}) = f(x^{(\nu)} - \alpha^{(\nu)} \nabla f(x^{(\nu)})).$$

- Even if we choose $\alpha^{(\nu)}$ at each iteration to minimize $f\left(x^{(\nu)} - \alpha^{(\nu)} \nabla f(x^{(\nu)})\right)$ *exactly* with respect to $\alpha^{(\nu)}$, it can take many iterations to find the minimum of a quadratic function having eccentric contour sets.
- The iterates will "zig-zag" back and forth across the axes of the eccentric contour sets, making slow progress towards $x^\star$.
- Non-quadratic functions with eccentric contour sets will exhibit similarly poor behavior using the steepest descent step direction.

Example, continued

- Using the function defined in (10.8), we obtain:

$$\forall x \in \mathbb{R}^2, \nabla f(x) = \begin{bmatrix} 2(x_1 - 1) - 1.8(x_2 - 3) \\ 2(x_2 - 3) - 1.8(x_1 - 1) \end{bmatrix}.$$

- Again, suppose that we use $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$ as the initial guess.
- Then:

$$\nabla f(x^{(0)}) = \begin{bmatrix} 2(3-1) - 1.8(-5-3) \\ 2(-5-3) - 1.8(3-1) \end{bmatrix} = \begin{bmatrix} 18.4 \\ -19.6 \end{bmatrix},$$

- and the steepest descent step direction at $x^{(0)}$ is $\Delta x^{(0)} = \begin{bmatrix} -18.4 \\ 19.6 \end{bmatrix}$.

- We update according to:

$$x^{(1)} = x^{(0)} + \alpha^{(0)}\Delta x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix} + \alpha^{(0)} \begin{bmatrix} -18.4 \\ 19.6 \end{bmatrix}.$$

- For the value of $\alpha^{(0)}$ that minimizes $f(x^{(0)} + \alpha^{(0)}\Delta x^{(0)})$ over choices of $\alpha^{(0)}$, $x^{(1)} \approx \begin{bmatrix} -1.8467 \\ 0.1628 \end{bmatrix}$, which is relatively far from the minimizer of f.
- Figure 10.9 illustrates the progress of iterations using steepest descent step direction, starting at $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, and assuming that at the ν-th iteration the step-size $\alpha^{(\nu)}$ is chosen to minimize $f(x^{(\nu)} + \alpha^{(\nu)}\Delta x^{(\nu)})$.
- Figure 10.9 shows that after two iterations of steepest descent we are close to the minimizer of this function.

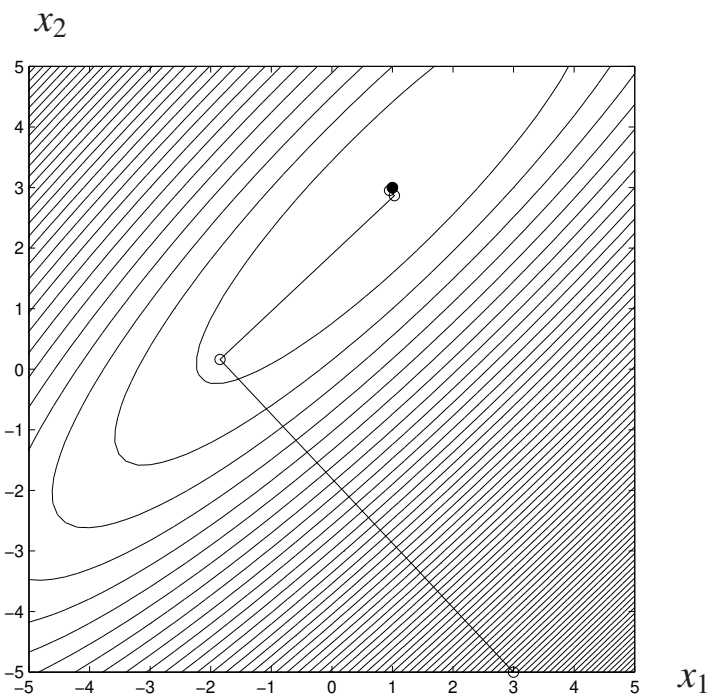

Fig. 10.9. Progress of iterations, shown as ○, using steepest descent step directions for an objective, defined in (10.8), with contour sets that are highly eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●. The initial guess was $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$.

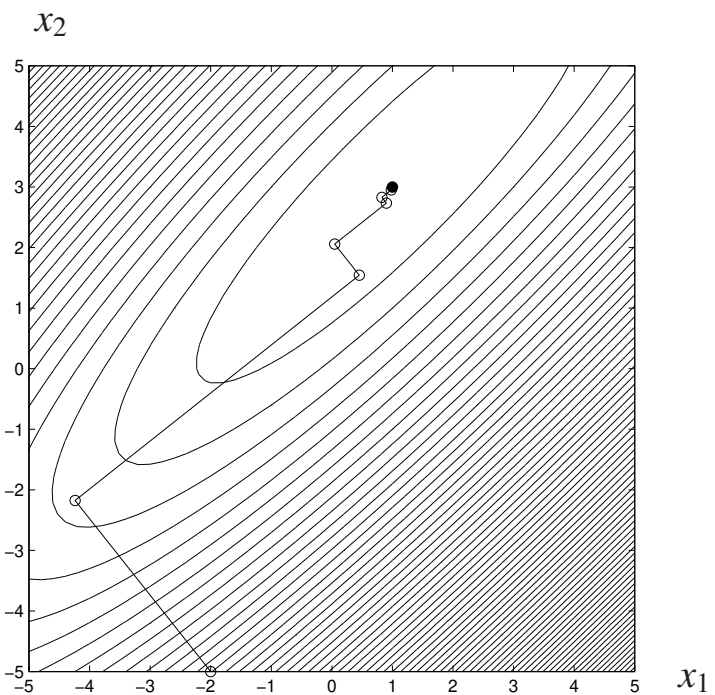

Fig. 10.10. Progress of iterations, shown as ○, using steepest descent step directions for an objective, defined in (10.8), with contour sets that are highly eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●. The initial guess was $x^{(0)} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$.

- However, starting at $x^{(0)} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$, the progress is much slower, as illustrated in Figure 10.10, requiring six steepest descent step directions to get close to the minimizer.
- In higher dimensions, with n larger than 2, the steepest descent algorithm will repeatedly take us in directions that do not point directly towards the minimizer.
- The steepest descent step direction can be arbitrarily close to being at *right angles* to the direction that points towards the minimizer.
- Moreover, we cannot expect to exactly minimize $f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)})$ over choices of $\alpha^{(\nu)}$ as assumed in Figures 10.9 and 10.10.
- This typically increases further the number of iterations required to find a useful answer.

10.2.1.5 Example with non-quadratic objective

$$\forall x \in \mathbb{R}^2, f(x) = 0.01 \times (x_1 - 1)^4 + 0.01 \times (x_2 - 3)^4 + (x_1 - 1)^2 + (x_2 - 3)^2 - 1.8(x_1 - 1)(x_2 - 3). \quad (10.9)$$

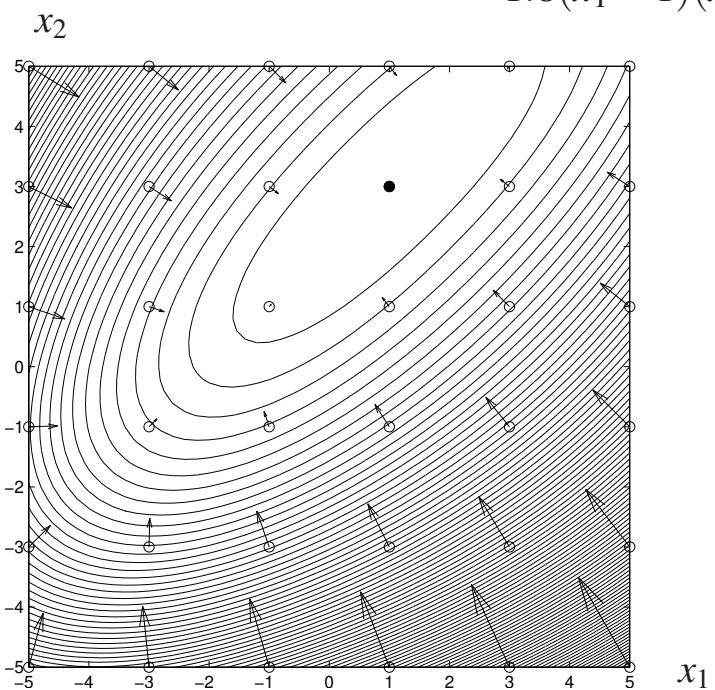

Fig. 10.11. Scaled versions of the steepest descent step directions for an objective, defined in (10.9), with contour sets that are perturbed eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●.

$$\forall x \in \mathbb{R}^2, \nabla f(x) = \begin{bmatrix} 0.04(x_1-1)^3 + 2(x_1-1) - 1.8(x_2-3) \\ 0.04(x_2-3)^3 - 1.8(x_1-1) + 2(x_2-3) \end{bmatrix}.$$

- Again, suppose that we use $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$ as the initial guess.
- Then, $\nabla f(x^{(0)}) = \begin{bmatrix} 18.72 \\ -40.08 \end{bmatrix}$ and the steepest descent step direction at $x^{(0)}$ is $\Delta x^{(0)} = \begin{bmatrix} -18.72 \\ 40.08 \end{bmatrix}$.
- We update according to:

$$x^{(1)} = x^{(0)} + \alpha^{(0)} \Delta x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix} + \alpha^{(0)} \begin{bmatrix} -18.72 \\ 40.08 \end{bmatrix}.$$

- Figure 10.12 shows the progress of a steepest descent algorithm assuming that at the ν-th iteration the step-size $\alpha^{(\nu)}$ is chosen to minimize $f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)})$.

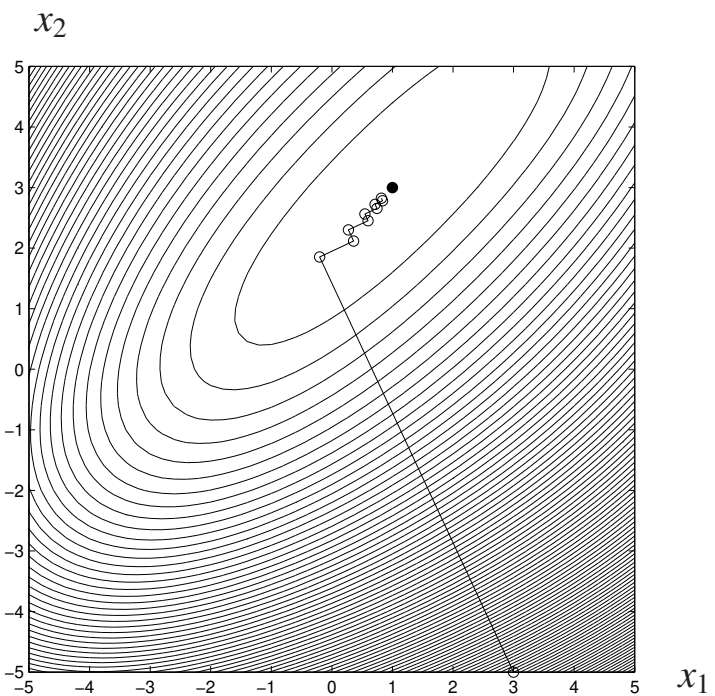

Fig. 10.12. Progress of iterations, shown as ○, using steepest descent step directions for an objective, defined in (10.9), with contour sets that are perturbed eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●. The initial guess was $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$.

Example with non-quadratic objective, continued

- Figure 10.13 shows the progress of a steepest descent algorithm starting at $x^{(0)} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$, again with the step-size chosen to minimize $f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)})$ at each iteration.
- The iterates again zig-zag back and forth across the axis of the contour sets and many iterations are required to approach the minimizer.

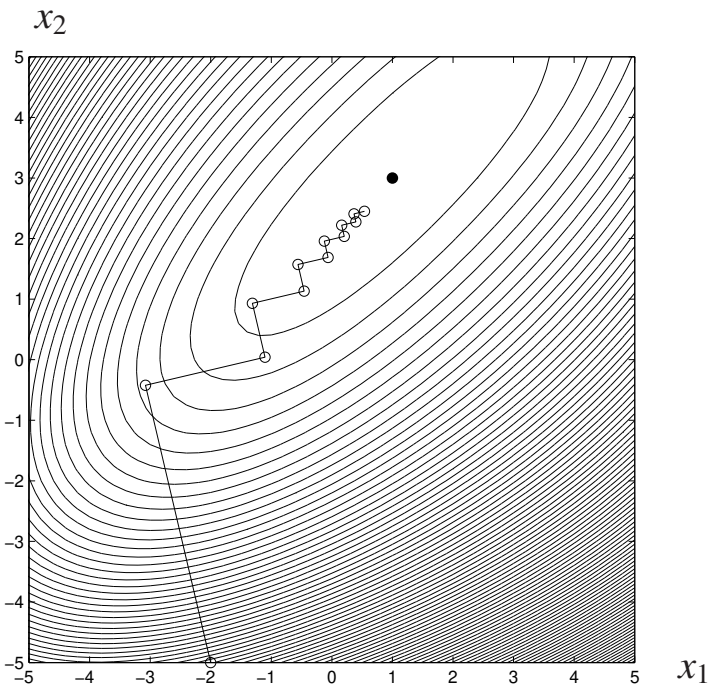

Fig. 10.13. Progress of iterations, shown as ○, using steepest descent step directions for an objective, defined in (10.9), with contour sets that are perturbed eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●. The initial guess was $x^{(0)} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$.

10.2.2 Solving $\nabla f(x) = \mathbf{0}$

- Another approach to minimizing f is based on the observation that $\nabla f(x) = \mathbf{0}$ is a system of either linear or non-linear equations having the same number of equations as variables.

10.2.2.1 Linear first-order necessary conditions

Analysis

- Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is quadratic of the form:

$$\forall x \in \mathbb{R}^n, f(x) = \frac{1}{2} x^\dagger Q x + c^\dagger x,$$

- In this case, the equations $\nabla f(x) = \mathbf{0}$ are linear and of the form $Qx + c = \mathbf{0}$.
- We can solve the equations:

$$Qx^\star = -c.$$

Example

$$
\begin{aligned}
\forall x \in \mathbb{R}^2, f(x) &= (x_1-1)^2+(x_2-3)^2-1.8(x_1-1)(x_2-3), \\
&= \frac{1}{2}x^\dagger Q x + c^\dagger x + \text{ constant}, \\
Q &= \nabla^2 f(x), \\
&= \begin{bmatrix} 2 & -1.8 \\ -1.8 & 2 \end{bmatrix}, \\
c &= \begin{bmatrix} 3.4 \\ -4.2 \end{bmatrix}.
\end{aligned}
$$

- Solving $Qx^\star = -c$ we obtain the minimizer $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Analysis

- Apply the Newton–Raphson update to solve $\nabla f(x) = \mathbf{0}$.

$$\begin{aligned} \nabla^2 f(x^{(\nu)}) \Delta x^{(\nu)} &= -\nabla f(x^{(\nu)}), \\ x^{(\nu+1)} &= x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)}, \end{aligned}$$

- The choice of step is called the **Newton–Raphson step direction** to minimize f.

Example with quadratic objective

- For a quadratic function, the necessary conditions are linear.
- Nevertheless, we can consider applying the Newton–Raphson update to solve them as though they were non-linear.
- For a quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ defined by:

$$\forall x \in \mathbb{R}^n, f(x) = \frac{1}{2} x^\dagger Q x + c^\dagger x,$$

- where $Q \in \mathbb{R}^{n \times n}$ and $c \in \mathbb{R}^n$, the Newton–Raphson step direction is the solution to $Q\Delta x^{(\nu)} = -Qx^{(\nu)} - c$.
- Using this update with step-size one yields a point satisfying the first-order necessary conditions for minimizing f.
- Figure 10.14 shows scaled versions of the Newton–Raphson step directions for the function (10.8) at various points.
- They all point towards the minimizer $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Example with quadratic objective, continued

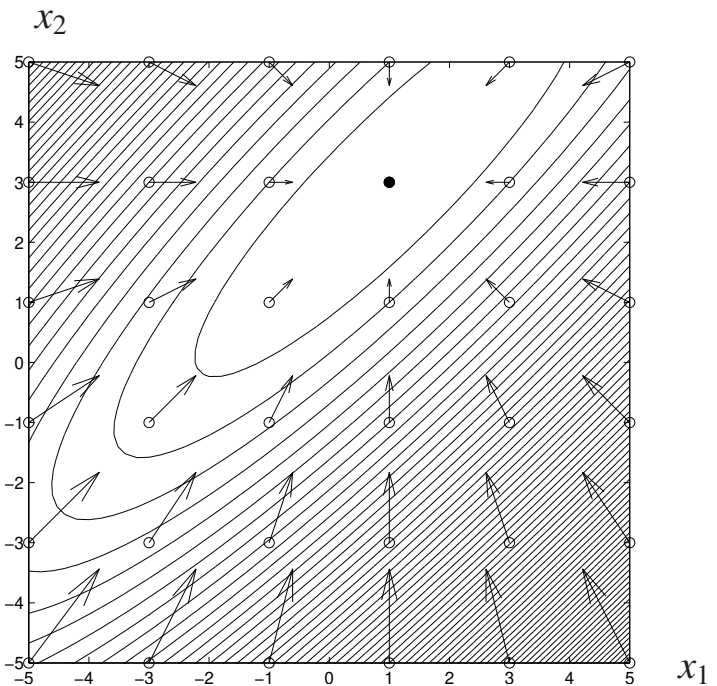

Fig. 10.14. Scaled versions of the Newton–Raphson step directions for an objective, defined in (10.8), with contour sets that are highly eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a $\bullet$.

Example with non-quadratic objective

$$\begin{aligned}\forall x \in \mathbb{R}^2, f(x) &= 0.01(x_1-1)^4 + 0.01(x_2-3)^4 + (x_1-1)^2 + (x_2-3)^2 \\ &\quad - 1.8(x_1-1)(x_2-3), \\ \forall x \in \mathbb{R}^2, \nabla^2 f(x) &= \begin{bmatrix} 0.12(x_1-1)^2+2 & -1.8 \\ -1.8 & 0.12(x_2-3)^2+2 \end{bmatrix}.\end{aligned}$$

- Again, suppose that we use $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$ as the initial guess.
- The Newton–Raphson step direction at $x^{(0)}$ is the solution to:

$$\begin{aligned}\begin{bmatrix} 2.48 & -1.8 \\ -1.8 & 9.68 \end{bmatrix} \Delta x^{(0)} &= \begin{bmatrix} -18.72 \\ 40.08 \end{bmatrix}, \\ \Delta x^{(0)} &\approx \begin{bmatrix} -5.250 \\ 3.164 \end{bmatrix}.\end{aligned}$$

Example with non-quadratic objective, continued

- We update according to:

$$x^{(1)} = x^{(0)} + \alpha^{(0)} \Delta x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix} + \alpha^{(0)} \begin{bmatrix} -5.250 \\ 3.164 \end{bmatrix}.$$

- For step-size $\alpha^{(0)} = 1$, we obtain $x^{(1)} = \begin{bmatrix} -2.250 \\ -1.836 \end{bmatrix}$.
- Figure 10.15 shows the progress of a Newton–Raphson algorithm starting at $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$ and assuming that at the ν-th iteration the step-size $\alpha^{(\nu)}$ were chosen to minimize $f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)})$.

Example with non-quadratic objective, continued

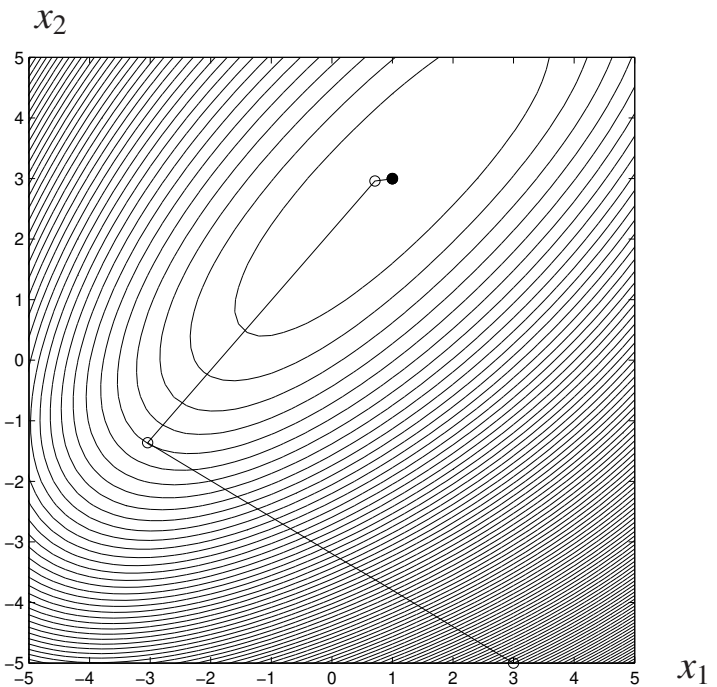

Fig. 10.15. Progress of iterations, shown as ○, using Newton–Raphson step directions for an objective, defined in (10.9), with contour sets that are perturbed eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●. The initial guess was $x^{(0)} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$.

Example with non-quadratic objective, continued

- Figure 10.16 shows the progress of a Newton–Raphson algorithm starting at $x^{(0)} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$, again with the step-size chosen to minimize $f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)})$ at each iteration.
- The progress is much faster than for the steepest descent step direction for the same value of initial guess.

Example with non-quadratic objective, continued

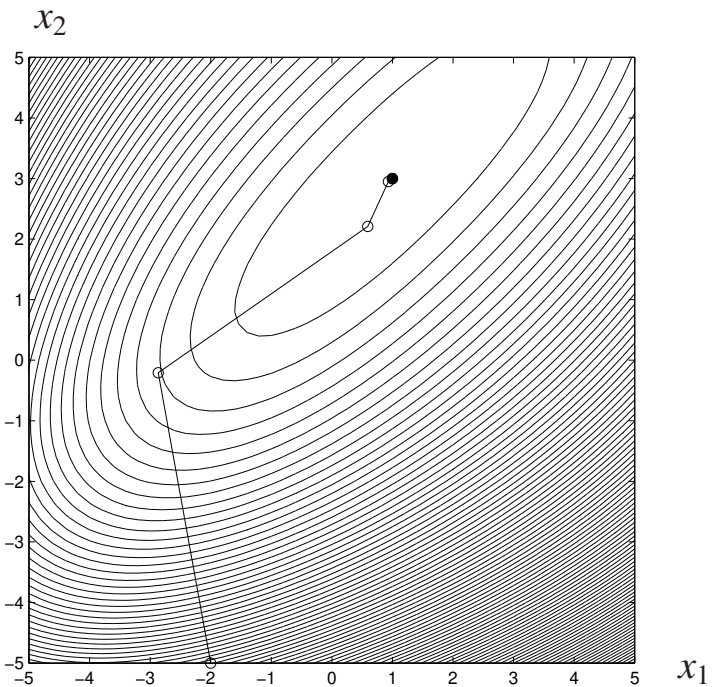

Fig. 10.16. Progress of iterations, shown as ○, using Newton–Raphson step directions for an objective, defined in (10.9), with contour sets that are perturbed eccentric ellipses. The contours of the function decrease towards $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is shown as a ●. The initial guess was $x^{(0)} = \begin{bmatrix} -2 \\ -5 \end{bmatrix}$.

10.2.2.3 Advantages

- Convergence to the solution of $\nabla f(x) = \mathbf{0}$ will be rapid, at least for initial guesses that are near to a solution of the equations or after the iterate becomes close to a solution of the equations.
- If f is quadratic then, as discussed in Section 10.2.2.2, the Newton–Raphson step direction with step-size $\alpha^{(\nu)} = 1$ takes us to a critical point in just one iteration.
- Since $\nabla^2 f(x)$ is symmetric, we can take advantage of symmetry in factorization.

10.2.2.4 Disadvantages

- For non-quadratic objectives and particularly at points that are far from the minimizer, the Newton–Raphson step direction is not necessarily a better direction than the steepest descent step direction.
- Factorization of the Hessian may require considerable effort if n is large or the Hessian is dense.
- If $\nabla f(x^{(\nu)})$ is not known analytically then it may be difficult or impossible to directly calculate $\nabla^2 f(x^{(\nu)})$.
- If $\nabla^2 f(x^{(\nu)})$ is not positive definite, then the Newton–Raphson update may take us towards a maximum or a point of inflection.

10.2.3 Generalization of Newton–Raphson and steepest descent

- In this section we generalize the Newton–Raphson and steepest descent updates in a way that can combine the advantages of each approach.

10.2.3.1 Uniform treatment of updates

$$\Delta x^{(\nu)} = -M \nabla f(x^{(\nu)}), \tag{10.10}$$

- with $M \in \mathbb{R}^{n \times n}$ positive definite as in Corollary 10.2 to guarantee descent.
- $M = \mathbf{I}$ yields the steepest descent step direction.
- $M = [\nabla^2 f(x^{(\nu)})]^{-1}$ (if the Hessian $\nabla^2 f$ is positive definite) yields the Newton–Raphson step direction.

10.2.3.2 Modified update

- To calculate $\Delta x^{(\nu)}$ satisfying (10.10), we would solve the linear system:

$$\nabla^2 f(x^{(\nu)})\Delta x^{(\nu)} = -\nabla f(x^{(\nu)}). \tag{10.11}$$

- Suppose that at the j-th stage of the factorization there are no positive diagonal pivots available.
- By Lemma 5.4, this means that $\nabla^2 f(x^{(\nu)})$ is not positive definite, so that the Newton–Raphson step direction, even if it is defined, may not be a descent direction.
- Let us modify the factorization by adding a positive quantity E_{jj} to $A^{(j)}_{jj}$ to make the pivot positive, where $A^{(j)}$ is the matrix obtained at the j-th stage of the factorization of $\nabla^2 f(x^{(\nu)})$.

Modified update, continued

- Adding E_{jj} to $A^{(j)}_{jj}$ is equivalent to adding the matrix:

$$E = \begin{bmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & E_{jj} & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \tag{10.12}$$

- to $\nabla^2 f(x^{(\nu)})$.
- By construction, $\nabla^2 f(x^{(\nu)}) + E$ is symmetric and positive definite.
- Its inverse $M = [\nabla^2 f(x^{(\nu)}) + E]^{-1}$ exists and is also symmetric and positive definite.
- By Corollary 10.2, the search direction defined by (10.10) using this M is a descent direction.
- This is called a **modified factorization**.

10.2.3.3 Further variations

- We have considerable flexibility to either:

  (i) construct positive definite approximations to $[\nabla^2 f(x)]^{-1}$, or
  (ii) approximately solve the equation:

$$\nabla^2 f(x)\Delta x = -\nabla f(x),$$

- in a way that guarantees that for the resulting Δx we have that $\Delta x = -M\nabla f(x)$ for some positive definite M.

10.2.4 Step-size

10.2.4.1 Need for step-size selection

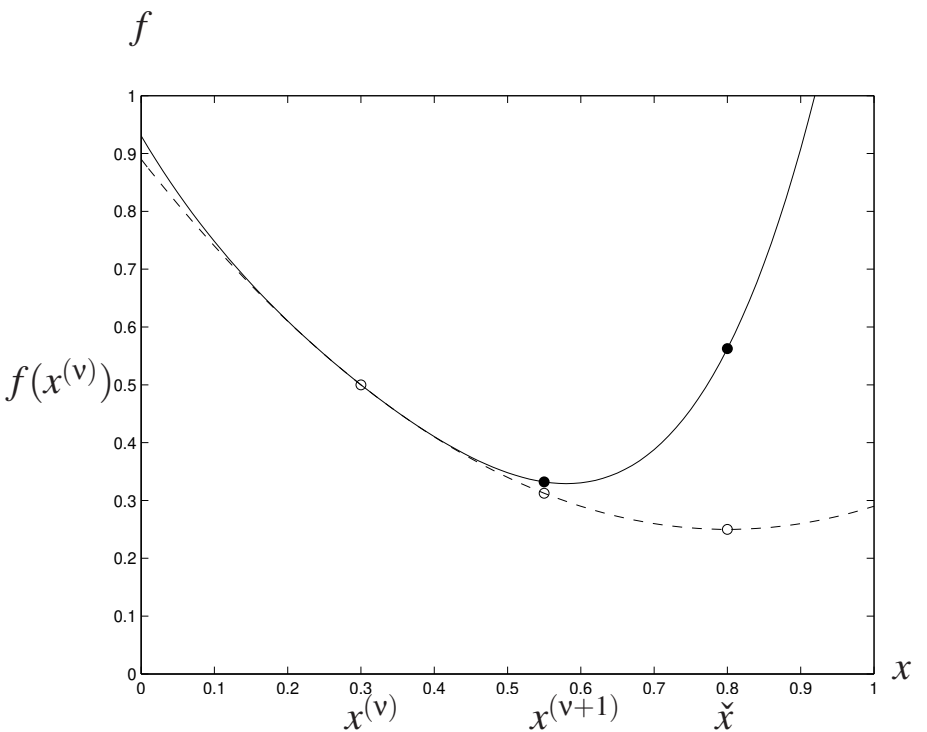

Fig. 10.17. The need for a step-size rule. The function f is illustrated with a solid line together with a quadratic approximation to it, illustrated as a dashed line. The quadratic approximation is a second-order Taylor approximation of f about $x^{(\nu)} = 0.3$.

Need for step-size selection, continued

- Suppose that we use the Newton–Raphson step direction to minimize the function shown in Figure 10.17, starting at $x^{(\nu)} = 0.3$.

$$\begin{aligned} \nabla^2 f(x^{(\nu)}) \Delta x^{(\nu)} &= -\nabla f(x^{(\nu)}), \\ \Delta x^{(\nu)} &= 0.5. \end{aligned}$$

- For this choice, $\check{x} = x^{(\nu)} + \Delta x^{(\nu)} = 0.8$ minimizes the quadratic approximation to f.
- However:

$$\begin{aligned} f(\check{x}) &= f(x^{(\nu)} + \Delta x^{(\nu)})), \\ &> f(x^{(\nu)}). \end{aligned}$$

- A step-size of $\alpha^{(\nu)} = 1$ would lead to an *increase* in the objective.

10.2.4.2 Armijo step-size rule

- Suppose that we had chosen $\alpha^{(\nu)}$ that is small enough so that f is accurately represented by a **second-order Taylor approximation** about $x^{(\nu)}$.
- Then:

$$
\begin{aligned}
& f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)}) \\
& \quad \approx \; f(x^{(\nu)}) + \alpha^{(\nu)} \left[\nabla f(x^{(\nu)})\right]^{\dagger} \Delta x^{(\nu)} + \frac{1}{2} (\alpha^{(\nu)})^2 [\Delta x^{(\nu)}]^{\dagger} \nabla^2 f(x^{(\nu)}) \Delta x^{(\nu)}, \\
& \qquad \text{by a second-order Taylor approximation,} \\
& \quad \approx \; f(x^{(\nu)}) + \alpha^{(\nu)} \left[\nabla f(x^{(\nu)})\right]^{\dagger} \Delta x^{(\nu)} - \frac{1}{2} (\alpha^{(\nu)})^2 [\Delta x^{(\nu)}]^{\dagger} \nabla f(x^{(\nu)}), \\
& \qquad \text{assuming that } \Delta x^{(\nu)} \text{ approximately solves } \nabla^2 f(x^{(\nu)}) \Delta x^{(\nu)} = -\nabla f(x^{(\nu)}), \\
& \quad = \; f(x^{(\nu)}) + \alpha^{(\nu)} \left(1 - \frac{1}{2} \alpha^{(\nu)}\right) \left[\nabla f(x^{(\nu)})\right]^{\dagger} \Delta x^{(\nu)}, \\
& \quad \leq \; f(x^{(\nu)}) + \frac{1}{2} \alpha^{(\nu)} \left[\nabla f(x^{(\nu)})\right]^{\dagger} \Delta x^{(\nu)}. \qquad (10.13)
\end{aligned}
$$

Armijo step-size rule, continued

- In practice, the reduction may not be as small as predicted by (10.13) and we may have to accept a smaller reduction.
- We choose an acceptance tolerance $0 < \delta < 1$.
- We start with tentative step-size $\alpha^{(\nu)} = 1$ and calculate the trial objective $f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)})$.
- The step-size is accepted if:

$$f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)}) \leq f(x^{(\nu)}) + \frac{\delta}{2} \alpha^{(\nu)} \left[\nabla f(x^{(\nu)})\right]^{\dagger} \Delta x^{(\nu)}. \tag{10.14}$$

- Otherwise, reduce the step-size by a factor of, say, one half and repeat the process until an iterate is produced that satisfies (10.14).

10.2.4.3 Wolfe condition

- The rule for reducing the step-size discussed in the last section does not check for "improvement" in the gradient ∇f.
- An alternative that makes use of gradient information rather than objective values is provided by the **Wolfe condition**:

$$\left| [\nabla f(x^{(\nu)} + \alpha^{(\nu)} \Delta x^{(\nu)})]^{\dagger} \Delta x^{(\nu)} \right| \leq \eta \left| [\nabla f(x^{(\nu)})]^{\dagger} \Delta x^{(\nu)} \right|. \qquad (10.15)$$

- The Wolfe condition ensures that the **directional derivative** in the direction $\Delta x^{(\nu)}$ evaluated at the next iterate, $[\nabla f(x^{(\nu+1)})]^{\dagger} \Delta x^{(\nu)}$, is small compared to the directional derivative in the direction $\Delta x^{(\nu)}$ at the current iterate, $[\nabla f(x^{(\nu)})]^{\dagger} \Delta x^{(\nu)}$.

10.2.4.4 Combined Armijo and Wolfe conditions

- The Wolfe condition (10.15) is often used in conjunction with the Armijo condition (10.14).
- The Armijo condition (10.14) ensures that the step-size is not so large as to invalidate the quadratic approximation of the objective.
- The Wolfe condition (10.15) ensures that the gradient of the objective is reduced sufficiently by the step.

10.2.4.5 Curve fitting

- If f is relatively easy to evaluate, then we can evaluate it at several points along the line $x^{(\nu)} + \alpha\Delta x^{(\nu)}$ for $0 \leq \alpha \leq 1$ and then fit a polynomial curve.
- We can minimize a quadratic function of α using the following:
  - (i) If the coefficient of $(\alpha)^2$ in the quadratic function is positive, then the minimum of the function occurs at the point $x^{(\nu)} + \alpha\Delta x^{(\nu)}$ for α such that the derivative of the quadratic function with respect to α is equal to zero. If this value of α lies outside the range $[0, 1]$ then the closest end-point should be selected.
  - (ii) If the coefficient of $(\alpha)^2$ in the quadratic function is negative, then the minimizer is one of the end-points $\alpha = 0$ or $\alpha = 1$.

10.2.4.6 Trust region

- In a **trust region approach** the selection of an appropriate search direction and step-size both explicitly consider the region over which a second-order Taylor approximation represents the function f accurately.

10.2.5 Stopping criteria

- A typical criterion is to require that $\left\|\nabla f(x^{(\nu)})\right\|$ and $\left\|\Delta x^{(\nu)}\right\|$ be sufficiently small.
- By Theorem 2.6, if f is convex then any minimizer $x^\star$ of $f(x)$ must satisfy:

$$\begin{aligned} f(x^\star) &\geq f(x^{(\nu)}) + [\nabla f(x^{(\nu)})]^\dagger (x^\star - x^{(\nu)}), \\ &\geq f(x^{(\nu)}) - \left|[\nabla f(x^{(\nu)})]^\dagger (x^\star - x^{(\nu)})\right|, \\ &\geq f(x^{(\nu)}) - \left\|\nabla f(x^{(\nu)})\right\| \left\|x^\star - x^{(\nu)}\right\|. \end{aligned} \tag{10.16}$$

- If we know an *a priori* bound on the minimizer, then we can bound $\left\|x^\star - x^{(\nu)}\right\|$ independently of $x^\star$ by some $\overline{\rho}$.
- We can ensure that $f(x^{(\nu)})$ is within ε_f of the value of the global minimum by iterating until $\left\|\nabla f(x^{(\nu)})\right\| \leq \varepsilon_f / \overline{\rho}$.

Stopping criteria, continued

- The stopping criterion is often implemented in practice as a slightly different *relative* criterion by testing if:

$$\left\| \nabla f(x^{(\nu)}) \right\| \leq \frac{\varepsilon_f}{\overline{\rho}} \left(1 + |f(x^{(\nu)})| \right).$$

10.2.6 Avoiding critical points that are not minimizers

- If, at some iteration ν, we find that $\nabla f(x^{(\nu)}) = \mathbf{0}$ then our basic algorithm cannot make further progress.
- If f is convex, then by Corollary 10.6, $x^{(\nu)}$ is a minimizer and $f(x^{(\nu)})$ is a minimum.
- If f is not convex, then we may be at a point of inflection or a local maximizer.

Avoiding critical points that are not minimizers, continued

- In Figure 10.18, the iterate $x^{(\nu)} = 0.5$ is a horizontal inflection point of the objective.

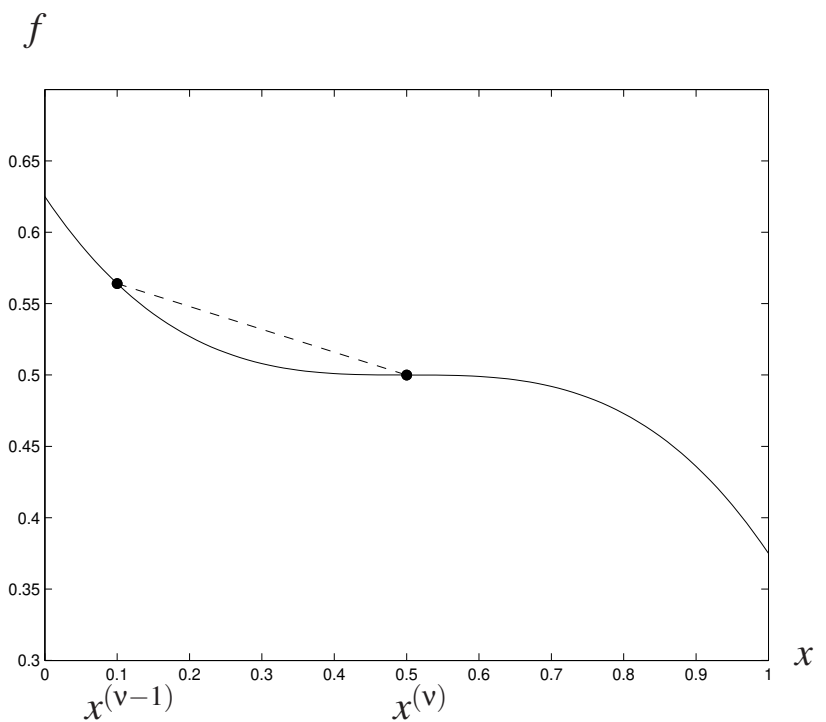

Fig. 10.18. Iterate that is a horizontal inflection point of the objective function.

Avoiding critical points that are not minimizers, continued

- If the first-order necessary conditions are satisfied, but we can detect that the current iterate is not a minimizer, then we can restart the algorithm by perturbing $x^{(\nu)}$ by a random amount to move it away from the point of inflection or local maximum.
- Alternatively, at a horizontal inflection, we can use the previous iterate in a secant approximation as discussed in Section 7.2.1.5, to seek a descent direction.
- For example, in Figure 10.18, using a secant approximation based on $x^{(\nu-1)}$ and $x^{(\nu)}$ would yield a descent direction.
- If we are not at a horizontal inflection point then another approach is to look for negative eigenvalues of the Hessian and move in the direction of the corresponding eigenvector.

10.3 Sensitivity

- Suppose that the objective f is *parameterized* by a parameter $\chi \in \mathbb{R}^s$. That is, $f : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$.
- We imagine that we have solved the unconstrained minimization problem:

$$\min_{x \in \mathbb{R}^n} f(x; \chi),$$

- for a base-case value of the parameters, say $\chi = \mathbf{0}$, to find the base-case minimizer $x^\star$.
- We now consider the sensitivity of the minimizer and minimum to variation of the parameters around $\chi = \mathbf{0}$.

Corollary 10.8 *Let $f : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}$ be twice partially differentiable with continuous second partial derivatives. Consider the minimization problem:*

$$\min_{x \in \mathbb{R}^n} f(x; \chi),$$

where $\chi \in \mathbb{R}^s$ is a parameter. Suppose that $x^\star$ is a local minimizer of this problem for the base-case value of the parameters $\chi = \mathbf{0}$. We call $x = x^\star$ a base-case minimizer. Define the (parameterized) Hessian $\nabla^2_{xx} f : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^{n \times n}$ by:

$$\forall x \in \mathbb{R}^n, \forall \chi \in \mathbb{R}^s, \nabla^2_{xx} f(x; \chi) = \frac{\partial^2 f}{\partial x^2}(x; \chi).$$

Suppose that $\nabla^2_{xx} f(x^\star; \mathbf{0})$ is positive definite, so that $x^\star$ satisfies the second-order sufficient conditions for the base-case problem. Then, there is a local minimizer of $f(x; \chi)$ for χ in a neighborhood of the base-case values of the parameters $\chi = \mathbf{0}$ and the local minimizer is a partially differentiable function of χ in this neighborhood. The sensitivity of the

local minimizer $x^\star$ with respect to variation of the parameters χ, evaluated at the base-case $\chi = \mathbf{0}$, is given by:

$$\frac{\partial x^\star}{\partial \chi}(\mathbf{0}) = -[\nabla^2_{xx} f(x^\star;\mathbf{0})]^{-1} K(x^\star;\mathbf{0}),$$

where $K : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^{n\times s}$ is defined by:

$$\forall x \in \mathbb{R}^n, \forall \chi \in \mathbb{R}^s, K(x;\chi) = \frac{\partial^2 f}{\partial x \partial \chi}(x;\chi).$$

The sensitivity of the corresponding local minimum $f^\star$ to variation of the parameters χ, evaluated at the base-case $\chi = \mathbf{0}$, is given by:

$$\frac{\partial f^\star}{\partial \chi}(\mathbf{0}) = \frac{\partial f}{\partial \chi}(x^\star;\mathbf{0}).$$

If $f(\bullet;\chi)$ is convex for each χ in a neighborhood of $\mathbf{0}$ then the minimizers and minima are global in this neighborhood.

Proof The sensitivity of the local minimizer follows from Corollary 7.5, noting that by assumption the Hessian is positive definite in a neighborhood of the base-case minimizer and parameters.
The sensitivity of the local minimum follows by totally differentiating the value of the local minimum $f^\star(\chi) = f(x^\star(\chi);\chi)$ with respect to χ. In particular,

$$\begin{aligned}
\frac{\partial f^\star}{\partial \chi}(\mathbf{0}) &= \frac{df(x^\star(\chi);\chi)}{d\chi}(\mathbf{0}), \\
&= \frac{\partial f}{\partial \chi}(x^\star;\mathbf{0}) + \frac{\partial f}{\partial x}(x^\star;\mathbf{0})\frac{\partial x^\star}{\partial \chi}(\mathbf{0}), \\
&\quad \text{on totally differentiating } f(x^\star(\chi);\chi) \text{ with respect to } \chi, \\
&= \frac{\partial f}{\partial \chi}(x^\star;\mathbf{0}),
\end{aligned}$$

since the first-order necessary conditions at the base-case are $\frac{\partial f}{\partial x}(x^\star;\mathbf{0}) = \mathbf{0}$.
The global results follow from Corollary 10.6. □

Discussion

- If $\nabla^2_{xx} f(x^\star; \mathbf{0})$ has already been factorized then each sensitivity of $x^\star$ with respect to an entry of χ requires only a forwards and backwards substitution.
- The sensitivity of the local minimum is called **the envelope theorem**.

10.3.2 Example

- Consider the parameterized objective function $f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ defined by:

$$\forall x \in \mathbb{R}^2, \forall \chi \in \mathbb{R}, f(x;\chi) = (x_1 - \exp(\chi))^2 + (x_2 - 3\exp(\chi))^2 + 5\chi.$$

- This is a parameterized version of the function defined in (10.1).
- For $\chi = 0$, the parameterized function is the same as the function defined in (10.1) and from the discussion in Section 10.1.1.2 we know that the base-case unconstrained minimizer is $x^\star = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.
- By Corollary 10.8, there is a minimizer of $f(\bullet;\chi)$ for χ in a neighborhood of the base-case value of the parameter $\chi = 0$ and the minimizer is a differentiable function of χ in this neighborhood.
- The sensitivity of the minimizer $x^\star$ with respect to variation of the parameter χ, evaluated at the base-case $\chi = 0$, is given by:

$$\frac{\partial x^\star}{\partial \chi}(0) = -[\nabla^2_{xx} f(x^\star;0)]^{-1} K(x^\star;0),$$

Example, continued

- where $\nabla^2_{xx} f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^{2\times 2}$ and $K : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^{2\times 1}$ are defined by:

$$\begin{aligned}
\forall x \in \mathbb{R}^2, \forall \chi \in \mathbb{R}, \nabla^2_{xx} f(x;\chi) &= \frac{\partial^2 f}{\partial x^2}(x;\chi), \\
&= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \\
\nabla^2_{xx} f(x^\star;0) &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \\
\forall x \in \mathbb{R}^2, \forall \chi \in \mathbb{R}, K(x;\chi) &= \frac{\partial^2 f}{\partial x \partial \chi}(x;\chi), \\
&= \begin{bmatrix} -2\exp(\chi) \\ -6\exp(\chi) \end{bmatrix}, \\
K(x^\star;0) &= \begin{bmatrix} -2 \\ -6 \end{bmatrix},
\end{aligned}$$

- where we observe that $\nabla^2_{xx} f(x^\star;0)$ is positive definite.

Example, continued

- The sensitivity of the minimizer $x^\star$ to variation of the parameter χ, evaluated at the base-case $\chi = 0$, is:

$$\begin{aligned}
\frac{\partial x^\star}{\partial \chi}(0) &= -[\nabla^2_{xx} f(x^\star;0)]^{-1} K(x^\star;0), \\
&= -\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} -2 \\ -6 \end{bmatrix}, \\
&= \begin{bmatrix} 1 \\ 3 \end{bmatrix}.
\end{aligned}$$

Example, continued

- The sensitivity of the minimum $f^\star$ to variation of the parameter χ, evaluated at the base-case $\chi = 0$, is given by:

$$\frac{\partial f^\star}{\partial \chi}(0) = \frac{\partial f}{\partial \chi}(x^\star; 0).$$

- We have that:

$$\frac{\partial f}{\partial \chi}(x;\chi) = 2(x_1 - \exp(\chi))(-\exp(\chi)) + 2(x_2 - 3\exp(\chi))(-3\exp(\chi)) + 5,$$

- and so the sensitivity is:

$$\frac{\partial f^\star}{\partial \chi}(0) = \frac{\partial f}{\partial \chi}(x^\star; 0) = 5.$$

10.4 Summary

- Descent directions,
- Optimality conditions,
- Algorithms,
- Sensitivity analysis.

11

Solution of the unconstrained minimization case studies

- Multi-variate linear regression case study in Section 11.1, and
- Power system state estimation case study in Section 11.2.

11.1 Multi-variate linear regression

11.1.1 Transformation of objective

- Recall Problem (9.7):

$$\max_{x\in\mathbb{R}^n} \phi(\zeta(1),\ldots,\zeta(m);\psi(1),\ldots,\psi(m),x),$$

- where $\phi:\mathbb{R}^n\to\mathbb{R}$ was defined in (9.6), which we repeat here:

$$\begin{aligned}\forall x\in\mathbb{R}^n, &\phi(\zeta(1),\ldots,\zeta(m);\psi(1),\ldots,\psi(m),x)\\ &= \prod_{\ell=1}^{m}\frac{1}{\sqrt{2\pi}\sigma_\ell}\exp\left(-\frac{(\psi(\ell)^\dagger\beta+\gamma-\zeta(\ell))^2}{2(\sigma_\ell)^2}\right).\end{aligned}$$

- First define $\hat{f}:\mathbb{R}^n\to\mathbb{R}$ by:

$$\forall x\in\mathbb{R}^n, \hat{f}(x) = -\ln(\phi(\zeta(1),\ldots,\zeta(m);\psi(1),\ldots,\psi(m),x)).$$

Transformation of objective, continued

- Then:

$$\begin{aligned}
\forall x \in \mathbb{R}^n, \hat{f}(x) &= -\ln\left(\prod_{\ell=1}^{m} \frac{1}{\sqrt{2\pi}\sigma_\ell} \exp\left(-\frac{(\psi(\ell)^\dagger \beta + \gamma - \zeta(\ell))^2}{2(\sigma_\ell)^2}\right)\right), \\
&= -\sum_{\ell=1}^{m}\left[\ln\left(\frac{1}{\sqrt{2\pi}\sigma_\ell}\right) - \frac{(\psi(\ell)^\dagger \beta + \gamma - \zeta(\ell))^2}{2(\sigma_\ell)^2}\right], \\
&= \sum_{\ell=1}^{m}\left[\frac{(\psi(\ell)^\dagger \beta + \gamma - \zeta(\ell))^2}{2(\sigma_\ell)^2}\right] - \sum_{\ell=1}^{m}\ln\left(\frac{1}{\sqrt{2\pi}\sigma_\ell}\right),
\end{aligned}$$

- where we recall that:

$$x = \begin{bmatrix} \beta \\ \gamma \end{bmatrix} \in \mathbb{R}^n.$$

Transformation of objective, continued

- Assuming that $\sigma_\ell = \sigma, \forall \ell = 1, \ldots, m$, we can define $f : \mathbb{R}^n \to \mathbb{R}$ by:

$$
\begin{aligned}
\forall x \in \mathbb{R}^n, f(x) &= \sigma^2 \left[\hat{f}(x) + \sum_{\ell=1}^{m} \ln \left(\frac{1}{\sqrt{2\pi}\sigma_\ell} \right) \right], \\
&= \frac{1}{2} \sum_{\ell=1}^{m} (\psi(\ell)^\dagger \beta + \gamma - \zeta(\ell))^2, \\
&= \frac{1}{2} \sum_{\ell=1}^{m} (A_\ell x - b_\ell)^2, \\
&\quad \text{where } A_\ell = \begin{bmatrix} \psi(\ell)^\dagger & 1 \end{bmatrix} \in \mathbb{R}^{1 \times n} \text{ and } b_\ell = \zeta(\ell) \in \mathbb{R}, \\
&= \frac{1}{2} (Ax - b)^\dagger (Ax - b), \\
&\quad \text{where } A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \in \mathbb{R}^{m \times n} \text{ and } b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m, \\
&= \frac{1}{2} \|Ax - b\|_2^2 .
\end{aligned}
$$

Transformation of objective, continued

- By Theorem 3.1, so long as either:

  (i) Problem (9.7) has a maximum or

  (ii) the problem:

$$\min_{x \in \mathbb{R}^n} f(x), \tag{11.1}$$

  has a minimum,

- then they both have the same set of optimizers.
- Problem (11.1) involves minimizing (half of) the sum of squares of linear functions of x and is called a **linear least-squares problem**.
- We refer to the corresponding specification of the affine function defined in (9.1) as a **least-squares fit** to the data.
- The necessary conditions for a minimum of Problem (11.1) are a set of linear simultaneous equations.

11.1.2 Comparison of objectives

- The necessary conditions for a minimum of Problem (11.1) are a set of linear simultaneous equations.
- In contrast, the necessary conditions for a maximum of Problem (9.7) are a set of non-linear simultaneous equations since ϕ is non-quadratic.

11.1.3 Derivatives of objective

$$\forall x \in \mathbb{R}^n, \nabla f(x) = A^\dagger (Ax - b), \tag{11.2}$$

$$\forall x \in \mathbb{R}^n, \nabla^2 f(x) = A^\dagger A. \tag{11.3}$$

11.1.4 Optimality conditions

- $\nabla^2 f(x)$ is positive semi-definite.
- Therefore the objective f is convex.
- First-order conditions are sufficient.
- Solving either Problem (11.1) or Problem (9.7) yields the same set of minimizers.
- In summary, by solving $\nabla f(x) = \mathbf{0}$ for $x^\star = \begin{bmatrix} \beta^\star \\ \gamma^\star \end{bmatrix}$ we will find a maximizer of Problem (9.7).
- Setting $\nabla f(x) = \mathbf{0}$ and re-arranging, we obtain:

$$\mathcal{A}x = \mathcal{B}, \tag{11.4}$$

- where $\mathcal{A} = A^\dagger A$ and $\mathcal{B} = A^\dagger b$.

11.1.5 Further transformation

- The condition number of $A^\dagger A$ can be large.
- Instead of calculating and factorizing $A^\dagger A$, we QR factorize A itself to obtain (ignoring any permutations of the rows or columns of A):

$$A = QR,$$

- with $Q \in \mathbb{R}^{m\times m}$ unitary, $R = \begin{bmatrix} U \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m\times n}$ upper triangular, with $U \in \mathbb{R}^{n\times n}$ upper triangular and U is non-singular if A has linearly independent columns.
- We have:

$$\begin{aligned} \forall x \in \mathbb{R}^n, f(x) &= \frac{1}{2}(Ax-b)^\dagger(Ax-b), \\ &= \frac{1}{2}(x^\dagger A^\dagger - b^\dagger)(Ax-b), \\ &= \frac{1}{2}(x^\dagger R^\dagger Q^\dagger - b^\dagger)(QRx-b), \text{ by definition of } QR, \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}(x^\dagger R^\dagger Q^\dagger - b^\dagger QQ^\dagger)(QRx - QQ^\dagger b), \text{ since } Q \text{ is unitary,} \\
&= \frac{1}{2}(x^\dagger R^\dagger - b^\dagger Q)Q^\dagger Q(Rx - Q^\dagger b), \text{ on factorizing,} \\
&= \frac{1}{2}(x^\dagger R^\dagger - b^\dagger Q)(Rx - Q^\dagger b), \text{ because } Q \text{ is unitary,} \\
&= \frac{1}{2}\left(x^\dagger \begin{bmatrix} U \\ \mathbf{0} \end{bmatrix}^\dagger - b^\dagger Q\right)\left(\begin{bmatrix} U \\ \mathbf{0} \end{bmatrix} x - Q^\dagger b\right), \text{ where } R = \begin{bmatrix} U \\ \mathbf{0} \end{bmatrix}, \\
&= \frac{1}{2}\left\| \begin{bmatrix} U \\ \mathbf{0} \end{bmatrix} x - Q^\dagger b \right\|_2^2, \\
&= \frac{1}{2}\left\| \begin{bmatrix} U \\ \mathbf{0} \end{bmatrix} x - \begin{bmatrix} [Q^{\|}]^\dagger \\ [Q^{\perp}]^\dagger \end{bmatrix} b \right\|_2^2, \text{ where } Q = \begin{bmatrix} Q^{\|} & Q^{\perp} \end{bmatrix}, \\
&\qquad \text{with } Q^{\|} \in \mathbb{R}^{m \times n'}, Q^{\perp} \in \mathbb{R}^{m \times (m-n')},
\end{aligned}
$$

$$= \frac{1}{2}\left\| \begin{bmatrix} Ux - [Q^{\|}]^{\dagger} b \\ \mathbf{0}x - [Q^{\perp}]^{\dagger} b \end{bmatrix} \right\|_2^2,$$

$$= \frac{1}{2}\left\| Ux - [Q^{\|}]^{\dagger} b \right\|_2^2 + \frac{1}{2}\left\| [Q^{\perp}]^{\dagger} b \right\|_2^2, \text{ by definition of the } L_2 \text{ norm.}$$

- Geometrically, we have resolved the vector $Ax - b$ into the sum of two vectors:
  $Ux - [Q^{\|}]^{\dagger} b$, which depends on x, and
  $\mathbf{0}x - [Q^{\perp}]^{\dagger} b = -[Q^{\perp}]^{\dagger} b$, which does not depend on x.
- The columns $Q^{\|}$ are such that $[Q^{\|}]^{\dagger} b$ "aligns" with Ux.
- The columns $Q^{\perp}$ are such that $(-[Q^{\perp}]^{\dagger} b)$ is perpendicular to Ux.
- If U is non-singular then the first-order necessary conditions for minimizing $\left\| Ux - [Q^{\|}]^{\dagger} b \right\|_2^2$ are $Ux = [Q^{\|}]^{\dagger} b$.

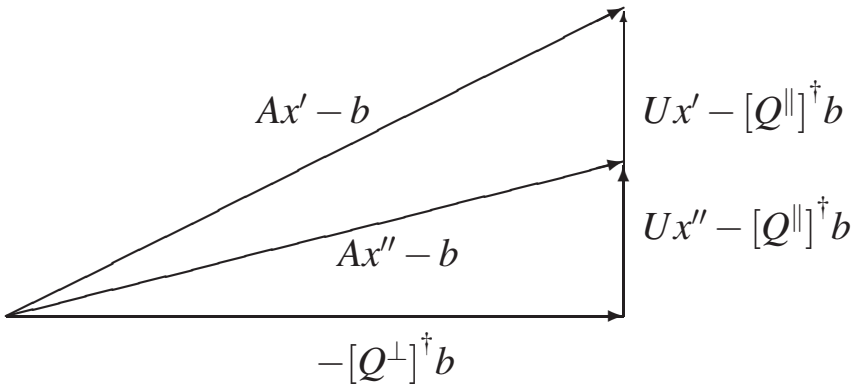

Fig. 11.1. Resolution of the vector $Ax - b$ into two perpendicular vectors for the values $x = x'$ and $x = x''$.

Further transformation, continued

- We can obtain the solution to Problem (11.1) by:
  - evaluating $y^\star = [Q^{\|}]^\dagger b$, and
  - performing a backwards substitution to solve $Ux^\star = y^\star$.
- The solution $x^\star = \begin{bmatrix} \beta^\star \\ \gamma^\star \end{bmatrix}$ specifies the maximum likelihood estimate of the relationship between the independent and dependent variables:

$$\forall \psi \in \mathbb{R}^{n-1}, \zeta = [\beta^\star]^\dagger \psi + \gamma^\star.$$

11.1.6 Relationship of optimality conditions to linear regression

- In designing the values of $\psi(\ell)$ for the trials, there are two related issues to be addressed:
  - (i) Providing enough variety in the trials to ensure that $\nabla^2 f = A^\dagger A$ is positive definite. We discuss this issue in Sections 11.1.6.1 and 11.1.6.2.
  - (ii) Providing enough redundancy so that the effects of measurement error can be "averaged out." We discuss this briefly in Section 11.1.6.3.

11.1.6.1 Insufficient variety in the trials

- If $\nabla^2 f(x)$ is singular then there will be many possible values of the parameters x that satisfy the maximum likelihood criterion in the model (9.1), based on the data from trials $\ell = 1, \ldots, m$.

11.1.6.2 Sufficient variety in the trials

- On the other hand, if there is an n element subset $\{\ell_1, \ell_2, \ldots, \ell_n\}$ of the trials $\{1, \ldots, m\}$ such that the n rows of A corresponding to these trials are linearly independent, then $\nabla^2 f(x) = A^\dagger A$ is non-singular.

11.1.6.3 Redundancy and validation of model

- We may want to find not only the maximum likelihood estimator but also estimate the variance of the error.
- In general, it requires that m be larger, and typically considerably larger, than n.

11.1.7 Changes in the problem

11.1.7.1 Additional trials

- If additional trials are added then there will be additional rows added to A and additional entries added to b, necessitating factorization of the augmented A.

11.1.7.2 Sensitivity

- We consider the sensitivity of the coefficients $\beta^\star$ and $\gamma^\star$ to changes in the measurements.
- That is, for each $\ell = 1, \ldots, m$, we will imagine that the ℓ-th measurement is actually $\zeta(\ell) + \chi_\ell$, with $\chi \in \mathbb{R}^m$.
- We calculate the sensitivity of $\beta^\star$ and $\gamma^\star$ to χ, evaluated at $\chi = \mathbf{0}$.
- By Corollary 10.8, the sensitivity of the minimizer $x^\star$ is given by:

$$\frac{\partial x^\star}{\partial \chi}(\mathbf{0}) = -\left[\nabla^2_{xx} f(x^\star; \mathbf{0})\right]^{-1} K(x^\star; \mathbf{0}),$$

- where $\nabla^2_{xx} f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n \times n}$ and $K : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n \times m}$ are defined by:

$$\begin{aligned} \forall x \in \mathbb{R}^n, \forall \chi \in \mathbb{R}^m, \nabla^2_{xx} f(x; \chi) &= \frac{\partial^2 f}{\partial x^2}(x; \chi), \\ &= A^\dagger A, \\ \forall x \in \mathbb{R}^n, \forall \chi \in \mathbb{R}^m, K(x; \chi) &= \frac{\partial^2 f}{\partial x \partial \chi}(x; \chi), \\ &= -A^\dagger. \end{aligned}$$

Sensitivity, continued

- That is, the sensitivity to χ_ℓ is given by $[A^\dagger A]^{-1} A^\dagger \mathbf{I}_\ell$, where $\mathbf{I}_\ell \in \mathbb{R}^m$ is a vector with zeros in all places except the ℓ-th place, which is a one.
- This is the same as the solution of a regression problem that had the same values of independent variables as in the base-case, but where the vector of measurements was changed from b to $\mathbf{I}_\ell$.
- Using the analysis in Section 11.1.5, we can calculate the sensitivity to χ_ℓ by:
  - evaluating $y = [Q^{\|}]^\dagger \mathbf{I}_\ell$, and
  - performing a backwards substitution to solve $U \dfrac{\partial x^\star}{\partial \chi}(0) = y$.

11.2 Power system state estimation

11.2.1 Transformation of objective

- We use a similar transformation to the one in Section 11.1.
- We define:

$$\forall x \in \mathbb{R}^n, f(x) = -\ln \phi(\tilde{G}; x) + \sum_{\ell \in \mathbb{M}} \ln \frac{1}{\sqrt{2\pi\sigma_\ell}}, \tag{11.5}$$

$$\begin{aligned} \forall x \in \mathbb{R}^n, f(x) &= \sum_{\ell \in \mathbb{M}} \frac{(\tilde{g}_\ell(x) - \tilde{G}_\ell)^2}{2\sigma_\ell^2}, \\ &= \frac{1}{2}(\tilde{g}(x) - \tilde{G})^\dagger [\Sigma]^{-2} (\tilde{g}(x) - \tilde{G}), \end{aligned} \tag{11.6}$$

- where:

  $\Sigma \in \mathbb{R}^{\mathbb{M} \times \mathbb{M}}$ is the diagonal matrix with ℓ-th diagonal entry equal to $\sigma_\ell, \ell \in \mathbb{M}$,

  $\tilde{g} : \mathbb{R}^n \to \mathbb{R}^{\mathbb{M}}$ is the vector of all measurement functions, and

  $\tilde{G} \in \mathbb{R}^{\mathbb{M}}$ is the vector of all measurements.

Transformation of objective, continued

- The transformed problem is:

$$\min_{x \in \mathbb{R}^n} f(x). \tag{11.7}$$

- We have a least-squares problem since the objective is the sum of squares of terms.
- Since each term $(\tilde{g}(x) - \tilde{G})$ is non-linear, we classify Problem (11.7) as a **non-linear least-squares problem**.

11.2.2 Derivatives of objective

$$\begin{aligned}
\forall x \in \mathbb{R}^n, \nabla f(x) &= \tilde{J}(x)^\dagger [\Sigma]^{-2} (\tilde{g}(x) - \tilde{G}), \\
&= \sum_{\ell \in \mathbb{M}} \nabla \tilde{g}_\ell(x) [\Sigma_\ell]^{-2} (\tilde{g}_\ell(x) - \tilde{G}_\ell), \qquad (11.8) \\
\forall x \in \mathbb{R}^n, \nabla^2 f(x) &= \tilde{J}(x)^\dagger [\Sigma]^{-2} \tilde{J}(x) + \sum_{\ell \in \mathbb{M}} \nabla^2 \tilde{g}_\ell(x) [\Sigma_\ell]^{-2} (\tilde{g}_\ell(x) - \tilde{G}_\ell), \\
& \qquad (11.9)
\end{aligned}$$

- where $\tilde{J}$ is the Jacobian of $\tilde{g}$ and $\nabla \tilde{g}_\ell$ is the transpose of the ℓ-th row of $\tilde{J}$.

11.2.3 Optimality conditions and algorithms

11.2.3.1 Qualitative comparison between Problems (9.8) and (11.7)

- The first-order necessary conditions for Problem (9.8), $\nabla\phi(\tilde{G};x) = \mathbf{0}$, are non-linear.
- The first-order necessary conditions for Problem (11.7), $\nabla f(x) = \mathbf{0}$, are also non-linear.
- Consider the measurement functions in detail:
  - (i) Each voltage magnitude measurement function, $\tilde{u}_k(x) = u_k$, is linear.
  - (ii) The real and reactive injection measurement functions and the real and reactive flow measurement functions are *approximately* linear. This observation and the expression for ∇f, (11.8), mean that the necessary conditions for Problem (11.7), $\nabla f(x) = \mathbf{0}$, are also approximately linear.
- The transformation (11.5) transforms a non-linear objective into an *approximately* quadratic objective.

Qualitative comparison between Problems (9.8) and (11.7), continued

- The necessary conditions for minimizing Problem (11.7) are *approximately* linear.
- We use the hypotheses of the chord and Kantorovich theorems to *qualitatively* compare the convergence properties of the Newton–Raphson update applied to:
  - Problem (9.8); that is, $\nabla\phi(x) = \mathbf{0}$, and
  - Problem (11.7); that is, $\nabla f(x) = \mathbf{0}$.
- Since ∇f is approximately linear, then $\nabla^2 f$ is approximately constant and a Lipschitz constant can be found for $\nabla^2 f$ that is smaller than a Lipschitz constant for $\nabla^2\phi$.
- We expect the radii ρ_-, ρ_+, and $\overline{\rho}$ defined in Theorems 7.3 and 7.4 to be larger for the problem of solving $\nabla f(x) = \mathbf{0}$ than for the problem of solving $\nabla\phi(x) = \mathbf{0}$.
- That is, we can expect to converge to a solution from a poorer initial guess if we apply the chord or Newton–Raphson methods to solve $\nabla f(x) = \mathbf{0}$ instead of applying it to solve $\nabla\phi(x) = \mathbf{0}$.

Hessian

- The Hessian $\nabla^2 f$ from (11.9) consists of the sum of two terms:
  - (i) $\tilde{J}(x)^\dagger [\Sigma]^{-2} \tilde{J}(x)$, which is of the form $A^\dagger A$ for $A = [\Sigma]^{-1} \tilde{J}(x)$ and so the matrix $\tilde{J}(x)^\dagger [\Sigma]^{-2} \tilde{J}(x)$, is positive semi-definite, and
  - (ii) $\sum_{\ell \in \mathbb{M}} \nabla^2 \tilde{g}_\ell(x) [\Sigma_\ell]^{-2} (\tilde{g}_\ell(x) - \tilde{G}_\ell)$, which can turn out to be not positive semi-definite.

Search direction

- Recall that in defining a search direction, we found that $\Delta x^{(\nu)} = -M \nabla f(x^{(\nu)})$ is a descent direction if M is positive definite.
- We know that $\tilde{J}(x)^\dagger [\Sigma]^{-2} \tilde{J}(x)$ is positive semi-definite, but we do not know if the Hessian is positive semi-definite.
- Instead of using the exact Newton–Raphson update, we approximate $\nabla^2 f$ by its first term:

$$\tilde{J}(x)^\dagger [\Sigma]^{-2} \tilde{J}(x). \tag{11.10}$$

Search direction, continued

- We solve for the approximate update direction:

$$\begin{aligned} \tilde{J}(x^{(\nu)})^{\dagger}[\Sigma]^{-2}\tilde{J}(x^{(\nu)})\Delta x^{(\nu)} &= -\nabla f(x^{(\nu)}), \\ &= \tilde{J}(x^{(\nu)})^{\dagger}[\Sigma]^{-2}(\tilde{G}-\tilde{g}(x^{(\nu)})). \end{aligned} \quad (11.11)$$

- This approximation is called the **Gauss–Newton method.**
- We must still consider the possibility that $\tilde{J}(x^{(\nu)})^{\dagger}[\Sigma]^{-2}\tilde{J}(x^{(\nu)})$ is not positive definite.
- We can follow the approach discussed in Section 10.2.3.2 and add terms to the diagonal of the matrix during factorization to ensure that the modified matrix is positive definite.

Search direction by solving a related linear least-squares problem

- The use of (11.11) to calculate a search direction suffers from a similar drawback to the solution of (11.4) in the linear case.
- By defining $A = [\Sigma]^{-1}\tilde{J}(x^{(\nu)})$ and $b = [\Sigma]^{-1}(\tilde{G} - \tilde{g}(x^{(\nu)}))$, note that (11.11) is equivalent to $A^{\dagger}A\Delta x^{(\nu)} = A^{\dagger}b$, which is the same form as the optimality condition for the multi-variate linear regression problem.
- We can therefore find $\Delta x^{(\nu)}$ by noting that $\Delta x^{(\nu)}$ is the solution to the *linear* least-squares problem:

$$\min_{\Delta x \in \mathbb{R}^n} \frac{1}{2} \|A\Delta x - b\|_2^2 . \tag{11.12}$$

Levenberg–Marquardt

- An alternative approach is to approximate the possibly not positive semi-definite term $\sum_{\ell \in \mathbb{M}} \nabla^2 \tilde{g}_\ell(x) [\Sigma_\ell]^{-2} (\tilde{g}_\ell(x) - \tilde{G}_\ell)$ by the positive definite matrix $\lambda \mathbf{I}$, where $\lambda > 0$ is chosen to be large enough to make the resulting approximation of the Hessian positive definite.
- This is called the **Levenberg–Marquardt method**. and is related to the trust region approach mentioned in Section 10.2.4.

Further approximation

- We can further approximate $\tilde{J}$ using the fast-decoupled or other approximations to the Jacobian of the power flow equations, as in the discussion of the solution of the power flow equations in Section 8.2.4.2.

11.2.4 Placement of meters in the system

11.2.4.1 Insufficient variety in the measurements

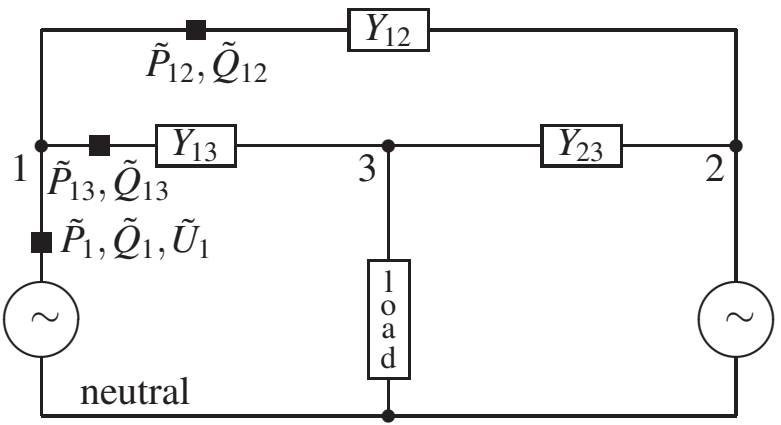

Fig. 11.2. The three-bus power system state estimation problem, repeated from Figure 9.2.

Insufficient variety in the measurements, continued

- If the measurements are not spread out throughout the system or if there is a measurement failure, then $\tilde{J}(x)^{\dagger}[\Sigma]^{-2}\tilde{J}(x)$ can be singular.
- For example, consider the system in Figure 9.2, which is repeated in Figure 11.2.
- The are five unknown variables: $u_1, \theta_2, u_2, \theta_3$, and u_3.
- There are seven measurements: $\tilde{P}_1, \tilde{Q}_1, \tilde{U}_1, \tilde{P}_{12}, \tilde{Q}_{12}, \tilde{P}_{13}$, and $\tilde{Q}_{13}$.
- However, since:

$$\begin{aligned} \tilde{p}_1(x) &= \tilde{p}_{12}(x) + \tilde{p}_{13}(x), \\ \tilde{q}_1(x) &= \tilde{q}_{12}(x) + \tilde{q}_{13}(x), \end{aligned}$$

- there is redundant information concerning bus 1.
- This would enable us to estimate the voltage magnitude and flows around node 1, even in the presence of measurement errors.
- There is just enough information to estimate all the voltage and flows in the system.

Insufficient variety in the measurements, continued

- Suppose that there is a failure of the voltage measurement in the system in Figure 11.2.
- In this case, there will be many sets of voltages and angles $\theta_2, |v_2|, \theta_3$, and $|v_3|$ that are consistent with maximizing the likelihood of the observed measurements.
- We say that the system is **unobservable**.
- If we are *designing* a measurement system, then singularity of $\tilde{J}(x)^{\dagger}[\Sigma]^{-2}\tilde{J}(x)$ for a candidate meter placement plan suggests that we should add more meters to the plan.
- If we are *operating* a measurement system and we find that because of, for example, meter failures, the matrix $\tilde{J}(x)^{\dagger}[\Sigma]^{-2}\tilde{J}(x)$ is singular, then we cannot estimate the state completely.
- In practice, in the latter case, the user of the software usually specifies **pseudo-measurements**; that is, guesses at what the actual measurement would be, based on experience, so that a rough estimate of the complete state can be found.

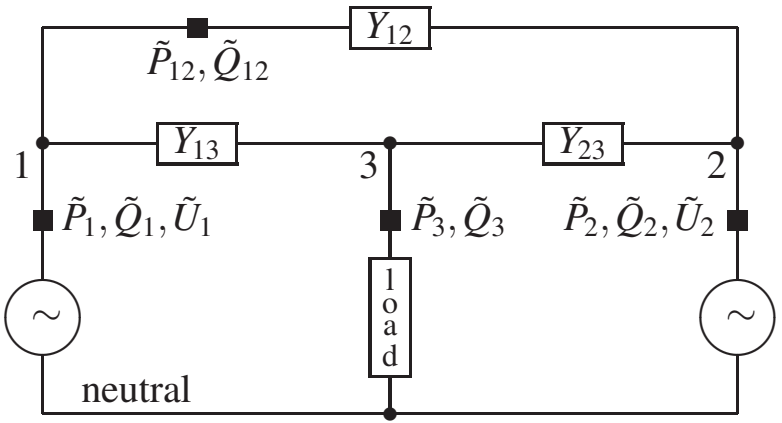

Fig. 11.3. The three-bus power system state estimation problem with spread out measurements repeated from Figure 9.3.

Sufficient variety in the measurements, continued

- Usually, if there is sufficient variety in the measurements, the positive semi-definite matrix $\tilde{J}(x)^{\dagger}[\Sigma]^{-2}\tilde{J}(x)$ will turn out to be positive definite for almost all, if not all values of x, and hence be non-singular.
- If it is non-singular then the approximate update equation (11.11) has a unique solution.
- For example, for the arrangement in Figure 9.3, which is repeated in Figure 11.3, for almost all values of x there is a five element subset of the rows of $\tilde{J}(x)$ that is linearly independent, so that $\tilde{J}(x)^{\dagger}[\Sigma]^{-2}\tilde{J}(x)$ is non-singular.
- This remains true even in the presence of a single failure of a voltage measurement.

11.2.4.3 Sensitivity

- We can consider variation of the estimate with variation in the measurement data.