と書き換えられる。

ただし，$\lambda_i = \dfrac{1}{n} - \overline{X}\omega_i$ とする。

λ_i は X の関数である。

● 例：X に観測誤差（measurement error）が含まれる場合：　真のモデルを

$$Y_i^* = \alpha + \beta X_i^*$$

とする。$(Y_i^*,\ X_i^*)$ は非確率変数とする。

しかし，$(Y_i^*,\ X_i^*)$ は観測されず，代わりに，$(Y_i,\ X_i)$ が観測されるものとする。

$(Y_i^*,\ X_i^*)$ と $(Y_i,\ X_i)$ との関係は以下の通りとする。

$$Y_i = Y_i^* + u_i, \qquad X_i = X_i^* + v_i$$

u_i，v_i は観測誤差と呼ばれるもので，

$$\mathrm{E}(u_i) = 0, \qquad \mathrm{V}(u_i) = \sigma_u^2$$
$$\mathrm{E}(v_i) = 0, \qquad \mathrm{V}(v_i) = \sigma_v^2$$

を仮定する。

さらに，u_i，v_i は互いに独立と仮定する。

すなわち，$i \neq j$ となるすべての i，j について $\mathrm{Cov}(u_i, u_j) = \mathrm{Cov}(v_i, v_j) = 0$，かつ，すべての i，j について $\mathrm{Cov}(u_i, v_j) = 0$ とする。

$Y_i^* = \alpha + \beta X_i^*$ に $Y_i = Y_i^* + u_i$，$X_i = X_i^* + v_i$ を代入する。

$$Y_i = \alpha + \beta X_i + (u_i - \beta v_i)$$

観測されるのは $(Y_i,\ X_i)$ なので，$(u_i - \beta v_i)$ を誤差項として，最小二乗法で $\hat{\beta}$ を求める。

まずは，X_i と $u_i - \beta v_i$ の共分散を求める（共分散がゼロかどうかを確認する）。

(* 復習) 共分散について：

2 つの確率変数 $(X,\ Y)$ を考える。

$\mathrm{E}(X) = \mu_x$，$\mathrm{E}(Y) = \mu_y$ とする。

共分散の定義は，$\mathrm{Cov}(X, Y) = \mathrm{E}\Big((X - \mu_x)(Y - \mu_y)\Big)$

書き換えると，

$\mathrm{Cov}(X, Y) = \mathrm{E}(XY - X\mu_y - \mu_x Y + \mu_x\mu_y) = \mathrm{E}(XY) - \mathrm{E}(X)\mu_y - \mu_x \mathrm{E}(Y) + \mu_x\mu_y = \mathrm{E}(XY) - \mu_x\mu_y$

となる。

この場合,

$$
\begin{aligned}
\mathrm{Cov}(X_i,\ u_i - \beta v_i) &= \mathrm{Cov}(X_i^* + v_i,\ u_i - \beta v_i) \\
&= \mathrm{E}\Big((X_i^* + v_i)(u_i - \beta v_i)\Big) - \mathrm{E}(X_i^* + v_i)\mathrm{E}(u_i - \beta v_i) \\
&= \mathrm{E}(X_i^* u_i + v_i u_i - X_i^* \beta v_i - \beta v_i^2) \\
&= \mathrm{E}(X_i^* u_i) + \mathrm{E}(v_i u_i) - \mathrm{E}(X_i^* \beta v_i) - \mathrm{E}(\beta v_i^2) \\
&= X_i^* \mathrm{E}(u_i) + \mathrm{E}(v_i u_i) - X_i^* \beta \mathrm{E}(v_i) - \beta \mathrm{E}(v_i^2) \\
&= -\beta \sigma_v^2 \neq 0
\end{aligned}
$$

となる。

したがって，観測できる $(Y_i,\ X_i)$ を用いて，β の最小二乗推定量 $\hat{\beta}$ は不偏推定量にはならない。

特に,

$$
\begin{aligned}
\hat{\beta} &= \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2} = \beta + \sum \omega_i (u_i - \beta v_i) \\
&= \beta + \frac{\sum (X_i - \overline{X})(u_i - \beta v_i)}{\sum (X_i - \overline{X})^2}
\end{aligned}
$$

$$= \beta + \frac{\frac{1}{n}\sum(X_i - \overline{X})(u_i - \beta v_i)}{\frac{1}{n}\sum(X_i - \overline{X})^2}$$

と書き換えられ，右辺第 2 項の分母は X_i の分散に対応し，分子は X_i と $(u_i - \beta v_i)$ との共分散 $-\beta\sigma_v^2$ に対応する。

　したがって，$n \longrightarrow \infty$ のとき，

$$\hat{\beta} \longrightarrow \beta + \frac{-\beta\sigma_v^2}{M_{xx}}$$

となる。右辺第 2 項の分母は必ず正，分子は β が正（負）の場合は負（正）となる。

　すなわち，

- $\beta > 0$ のとき，$\hat{\beta} \longrightarrow \beta - \dfrac{\beta\sigma_v^2}{M_{xx}} < \beta$
- $\beta < 0$ のとき，$\hat{\beta} \longrightarrow \beta - \dfrac{\beta\sigma_v^2}{M_{xx}} > \beta$

となる。

● X と u_i に相関がある場合の対処法：　$Y_i = \alpha + \beta X_i + u_i$ について，$\mathrm{Cov}(X_i, u_i) \neq 0$ のときを考える。

$$\hat{\beta} = \frac{\sum_i (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_i (X_i - \overline{X})^2} = \beta + \frac{\sum_i (X_i - \overline{X})u_i}{\sum_j (X_j - \overline{X})^2}$$

$$= \beta + \frac{\frac{1}{n}\sum_i (X_i - \overline{X})u_i}{\frac{1}{n}\sum_j (X_j - \overline{X})^2} \longrightarrow \beta + \frac{M_{xu}}{M_{xx}} \neq \beta$$

右辺第 2 項の分母は X_i の分散に相当し，分子は X_i と u_i の共分散に相当する（$n \longrightarrow \infty$ のときは，分子・分母を別々に計算することができる）。

$\mathrm{Cov}(X_i, u_i) \neq 0$ が問題となって，$\mathrm{E}(\hat{\beta}) \neq \beta$ となる。

よって，第 2 項の分子がゼロになるような修正を加えればよい。

$\mathrm{Cov}(Z_i, u_i) = 0$ となる Z_i が存在するとする。
このとき，下記のような推定量

$$\widetilde{\beta} = \frac{\sum_i (Z_i - \overline{Z})(Y_i - \overline{Y})}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})}$$

を考えてみよう。ただし，$\overline{Z} = \dfrac{1}{n} \displaystyle\sum_i Z_i$ とする。

(*) Z_i を X_i で置き換えると，$\widetilde{\beta}$ は最小二乗推定量 $\hat{\beta} = \dfrac{\sum_i (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_i (X_i - \overline{X})^2}$ に等しくなる。

$\widetilde{\beta}$ を変形していく。

$$
\begin{aligned}
\widetilde{\beta} &= \frac{\sum_i (Z_i - \overline{Z})(Y_i - \overline{Y})}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \\
&= \frac{\sum_i (Z_i - \overline{Z})Y_i - \overline{Y}\sum_i (Z_i - \overline{Z})}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \\
&= \frac{\sum_i (Z_i - \overline{Z})Y_i}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \qquad \left(= \sum_i \omega_i^* Y_i,\quad \text{ただし，}\ \omega_i^* = \frac{Z_i - \overline{Z}}{\sum_j (Z_j - \overline{Z})(X_j - \overline{X})}\right) \\
&= \frac{\sum_i (Z_i - \overline{Z})(\alpha + \beta X_i + u_i)}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \\
&= \frac{\alpha\sum_i (Z_i - \overline{Z}) + \beta\sum_i (Z_i - \overline{Z})X_i + \sum_i (Z_i - \overline{Z})u_i}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \\
&= \beta + \frac{\sum_i (Z_i - \overline{Z})u_i}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \qquad \left(= \beta + \sum_i \omega_i^* u_i\right) \\
&= \beta + \frac{\frac{1}{n}\sum_i (Z_i - \overline{Z})u_i}{\frac{1}{n}\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \longrightarrow \beta + \frac{0}{M_{zx}} = \beta
\end{aligned}
$$

2 行目の右辺の分子の第 2 項目は $\sum_i (Z_i - \overline{Z}) = \sum_i Z_i - n\overline{Z} = 0$ に注意。

3 行目では，$\widetilde{\beta} = \dfrac{\sum_i (Z_i - \overline{Z})Y_i}{\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} = \displaystyle\sum_i \omega_i^* Y_i$ と書き換えることができ，$\widetilde{\beta}$ も Y_i の線形推定量と言える。

ただし，$\omega_i^* = \dfrac{Z_i - \overline{Z}}{\sum_j (Z_j - \overline{Z})(X_j - \overline{X})}$ である（分母の添字を i から j に変更）。

4 行目の右辺分子の第 1 項目はゼロ，第 2 項目は $\sum_i (Z_i - \overline{Z})X_i = \sum_i (Z_i - \overline{Z})X_i - \sum_i (Z_i - \overline{Z})\overline{X} = \sum_i (Z_i - \overline{Z})(X_i - \overline{X})$ となるので分母と同じになる。このようにして，5 行目が得られる。

6 行目右辺第 2 項の分子は Z_i と u_i の共分散に対応し，分母は Z_i と X_i の共分散に対応し，n を大きくするとそれぞれゼロ，M_{zx} に収束するものとする。

すなわち，$\widetilde{\beta}$ は β の一致推定量となる。

n が大きければ，$\mathrm{E}(\omega_i^* u_i) \longrightarrow 0$ となる（分子・分母を別々に計算することができる）。

しかし，一般的には，$\mathrm{E}(\omega_i^* u_i) \neq 0$ なので（ω_i^* の分母は X_i に依存していて，X_i と u_i は共分散がゼロでないと仮定），$\mathrm{E}(\widetilde{\beta}) \neq \beta$ となり，$\widetilde{\beta}$ は不偏推定量にはならない。

Z_i を操作変数（instrumental variable）と呼ぶ。操作変数を用いた推定方法を操作変数法という。

Z_i の選択について，(i) Z_i と u_i は相関がない，(ii) Z_i と X_i は強い相関がある，という 2 つの条件が必要になる。

(ii) については，Z_i はもともと X_i の代わりに使うものなので，X_i と相関の強い Z_i が望ましい。

● Z_i **の選択について（その 1）：**　i が時間を表す場合（時系列データの場合），$Z_i = X_{i-1}$ を用いることが可能である。

$\mathrm{Cov}(X_i, u_i) \neq 0$ としても，X を一期ずらして $\mathrm{Cov}(X_{i-1}, u_i) = 0$ となるのは不自然ではない。

● Z_i **の選択について（その 2）：**　X_i の予測値 $\hat{X}_i$ を X_i の代わりに用いる。

e_i を誤差項として，

$$X_i = \gamma_0 + \gamma_1 W_{1i} + \gamma_2 W_{2i} + \cdots + \gamma_m W_{mi} + e_i$$

を最小二乗法で推定して，$\hat{\gamma}_0$，$\hat{\gamma}_1$，$\cdots$，$\hat{\gamma}_m$ を求める。

ただし，W_{1i}，W_{2i}，$\cdots$，W_{mi} は u_i と相関のない変数でなければならない。

W_{1i}，W_{2i}，$\cdots$，W_{mi} には，X_{i-1}，X_{i-2}，$\cdots$ のように X_i のラグ変数を用いてもよい。理由は，前述の通りで，X_i と u_i に相関があったとしても，X_i のラグ変数 X_{i-1}，X_{i-2}，$\cdots$ と u_i とに相関があるとは考えにくいからである。

X_i の予測値 $\hat{X}_i$ を求める。

$$\hat{X}_i = \hat{\gamma}_0 + \hat{\gamma}_1 W_{1i} + \hat{\gamma}_2 W_{2i} + \ \cdots \ + \hat{\gamma}_m W_{mi}$$

を Z_i として用いる。

Z_i は u_i と相関のない変数でなければならない。

$\hat{\gamma}_j$ は γ_j の一致推定量なので，$n \longrightarrow \infty$ のとき，

$$\hat{X}_i = \hat{\gamma}_0 + \hat{\gamma}_1 W_{1i} + \hat{\gamma}_2 W_{2i} + \ \cdots \ + \hat{\gamma}_m W_{mi} \longrightarrow \gamma_0 + \gamma_1 W_{1i} + \gamma_2 W_{2i} + \ \cdots \ + \gamma_m W_{mi}$$

となる。

操作変数法による推定量 $\widetilde{\beta}$ が β の一致推定量になる理由は,

$$\widetilde{\beta} = \beta + \frac{\frac{1}{n}\sum_i (Z_i - \overline{Z})u_i}{\frac{1}{n}\sum_i (Z_i - \overline{Z})(X_i - \overline{X})} \longrightarrow \beta + \frac{0}{M_{zx}} = \beta$$

から，操作変数 Z_i と誤差項 u_i との相関がゼロという条件（2 項目の分子）が重要なポイントとなっている。

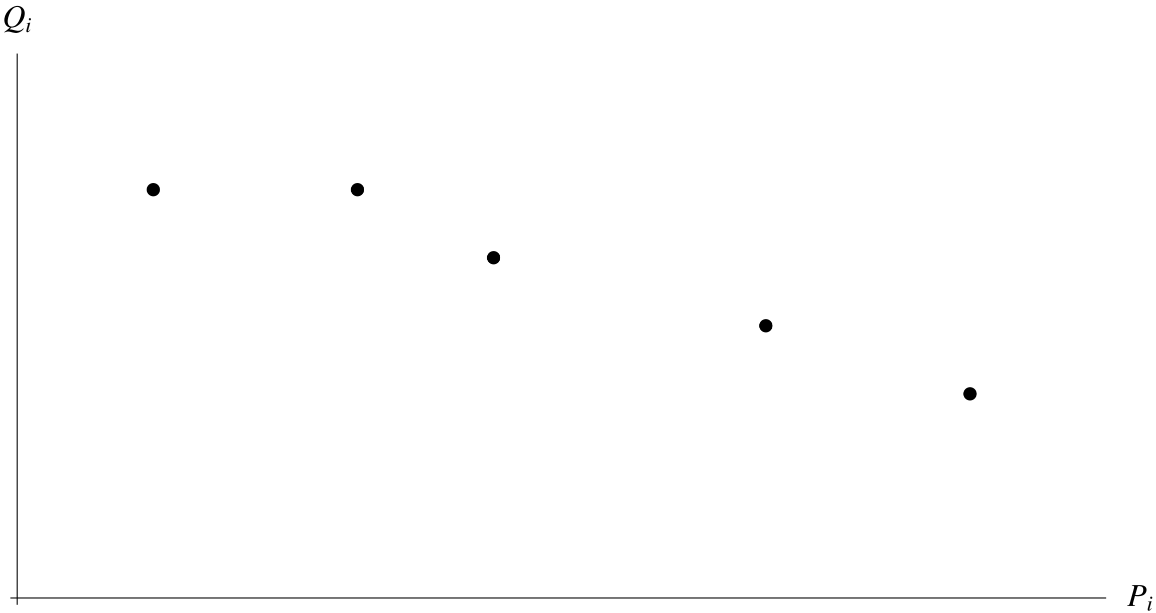
Z_i に $\hat{X}_i$ を用いると，$n \longrightarrow \infty$ のとき，$\hat{X}_i \longrightarrow \gamma_0+\gamma_1 W_{1i}+\gamma_2 W_{2i}+\cdots+\gamma_m W_{mi}$ となることから，W_{1i}，W_{2i}，$\cdots$，W_{mi} が u_i と相関がなければ，$\frac{1}{n}\sum_i(\hat{X}_i-\overline{X})u_i \longrightarrow 0$ となる（$\overline{\hat{X}}=\overline{X}$ に注意）。

この方法は，二段階最小二乗法（two-stage least squares method）と呼ばれる。1 段階目で $\hat{X}_i$ を求める。2 段階目で $\widetilde{\beta}$ を得る。

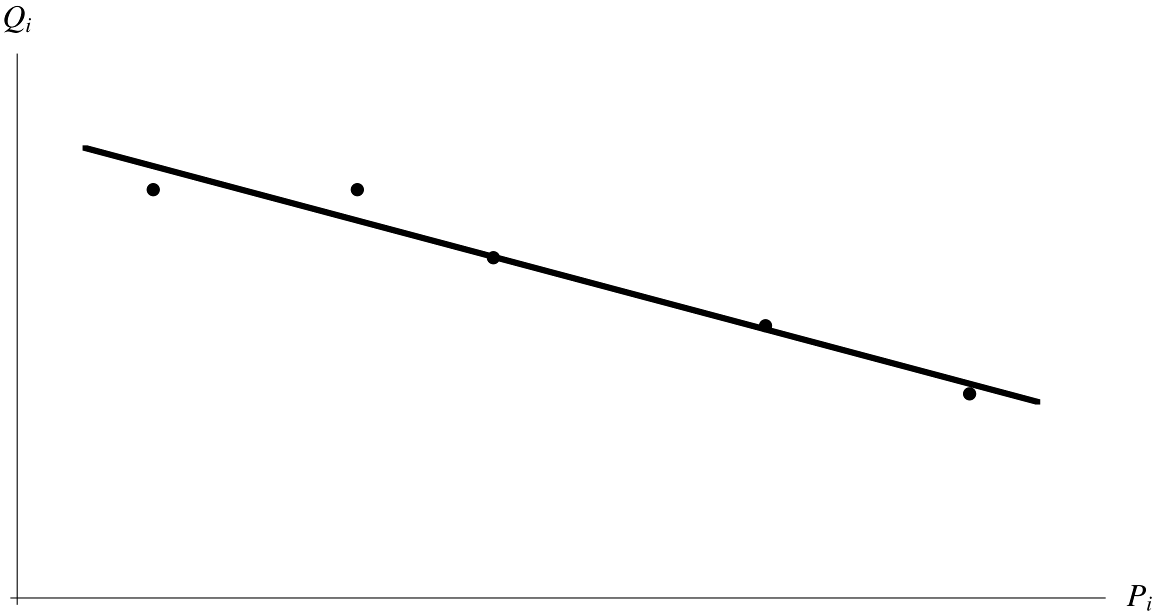
第13章　識別性について

13.1　例：需要関数・供給関数

ある財の需要関数・供給関数を推定することを考える。
需要量と供給量が等しいところで取引される。
すなわち，需要量と供給量が等しいところでデータが観測されるものとする。
下図のように黒丸で需要量 Q_i と価格 P_i のデータが観測されたとする。



最小二乗法によって，下図のように $(Q_i,\ P_i)$ の関係を表す線を引く。



この直線は何を表す直線か？

価格 P_i が上がれば，需要量 Q_i が下がるので，需要関数か？

それとも，供給関数か？

実は，この直線は需要関数でも供給関数でもどちらでもない。

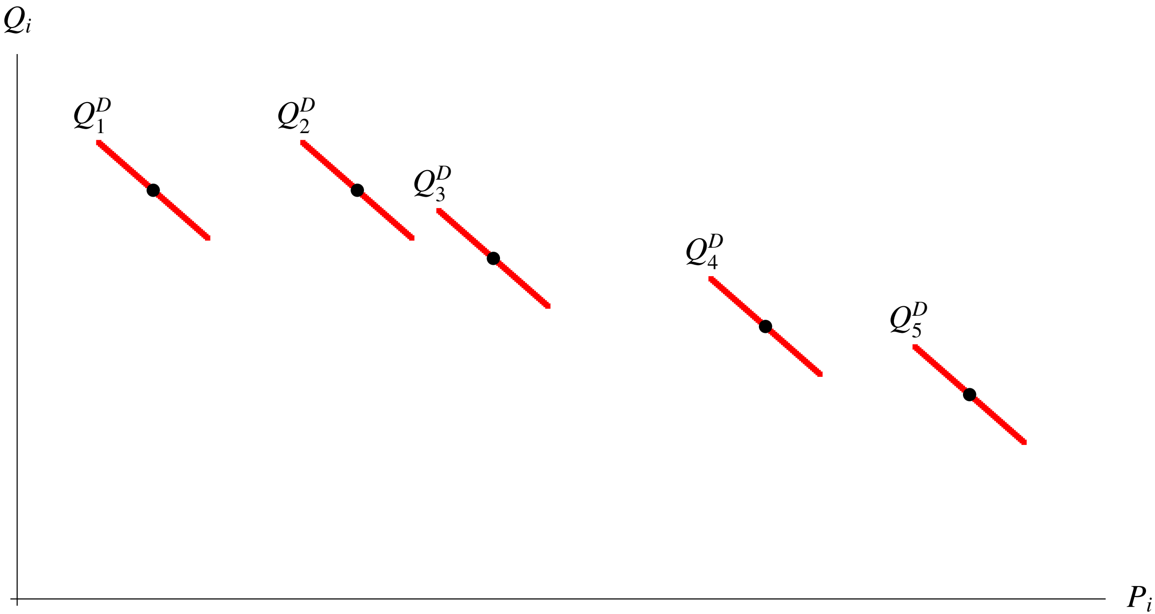
需要関数は，

$$Q_i^D = \alpha_0 + \alpha_1 P_i + \alpha_2 Y_i + u_i$$

となる。

価格 P_i が上がるにつれて，需要量 Q_i^D は減少する（右下がり）。すなわち，$\alpha_1 < 0$

所得 Y_i が増加するにつれて，需要量 Q_i^D が増加する。すなわち，$\alpha_2 > 0$



この例では，Q_1^D，Q_2^D，$\cdots$，Q_5^D の順に所得が増加する（Y_1 が最も小さく，Y_5 が最も大きい）。

所得 Y_i が大きくなるにつれて，切片が大きくなる（この場合，$\alpha_0 + \alpha_2 Y_i$ が切片）。

同じ価格の下では，所得が増加すれば需要量が増える。

供給関数は,

$$Q_i^S = \beta_0 + \beta_1 P_i + \beta_2 Z_i + v_i$$

と表される。

価格 P_i が上がるにつれて，供給量 Q_i^S は増加する（右上がり）。すなわち，$\beta_1 > 0$

価格が上がれば，生産を増やして売る方が儲けが多い。

変数 Z_i は財の種類によって異なる。

β_2 の符号も Z_i によって正負のどちらにもなり得る。

農作物の需給を考えるのであれば，供給量は天候に大きく依存するので，Z_i は日照時間，降水量などが適切な変数となる。

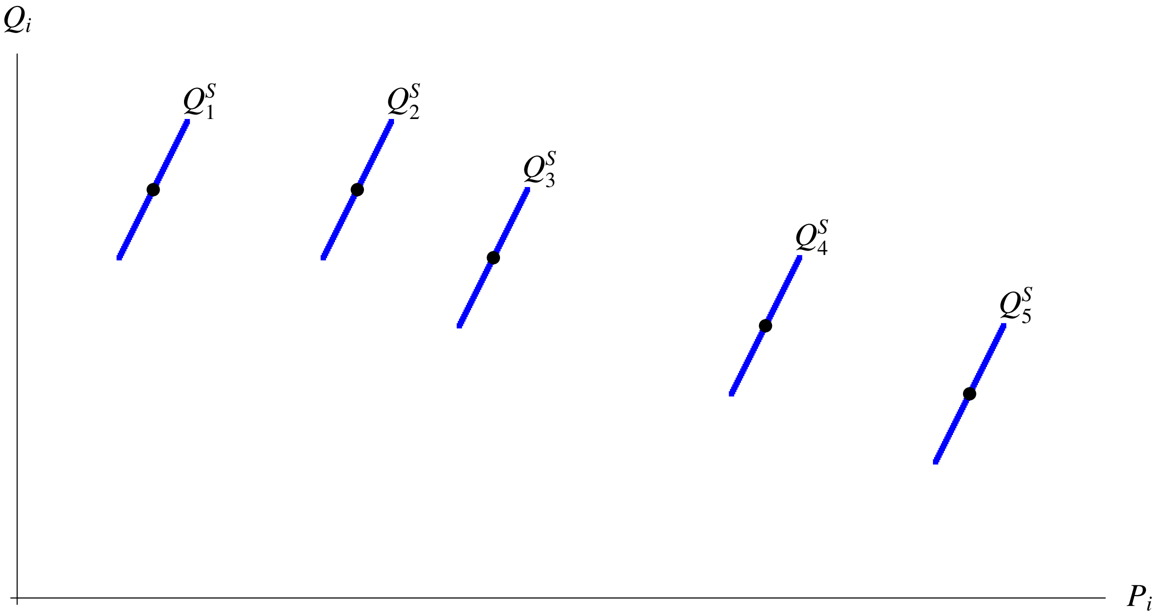
または，Z_i に作付面積なども考えられる。

作付面積が増えれば供給量も増える。

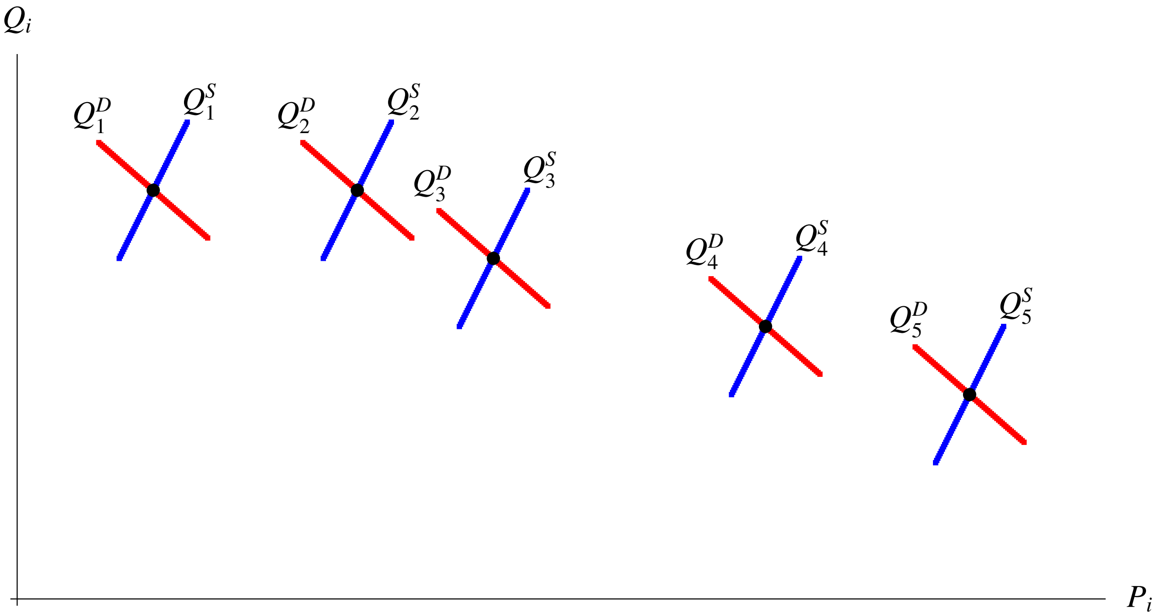
車，PC などは工場設備などの稼働率が考えられる。

いずれにしても，Z_i は供給関数にとって，重要な変数となる。

供給関数が次の図で表される。供給関数は右上がりで，$\beta_0 + \beta_2 Z_i$ が切片となる。



下図のように，需要と供給が交わるところで，需要量（供給量）と価格が決まる。



需要関数の形状が決まるためには，Q_i，P_i 以外の需要要因の変数が必要となる（この場合は，所得 Y_i）。

同様に，供給関数の形状が決まるためには，Q_i，P_i 以外の供給要因の変数が必要となる（この場合は，Z_i）。

この問題を，識別問題（identification problem）と呼ぶ。

需要関数と供給関数を区別して推定できるのかという問題である。

Q_i，P_i だけでは，何を推定しているのか分からなくなる。

何かを推定するうえでは，常に念頭に置いておかなければならない事項である。